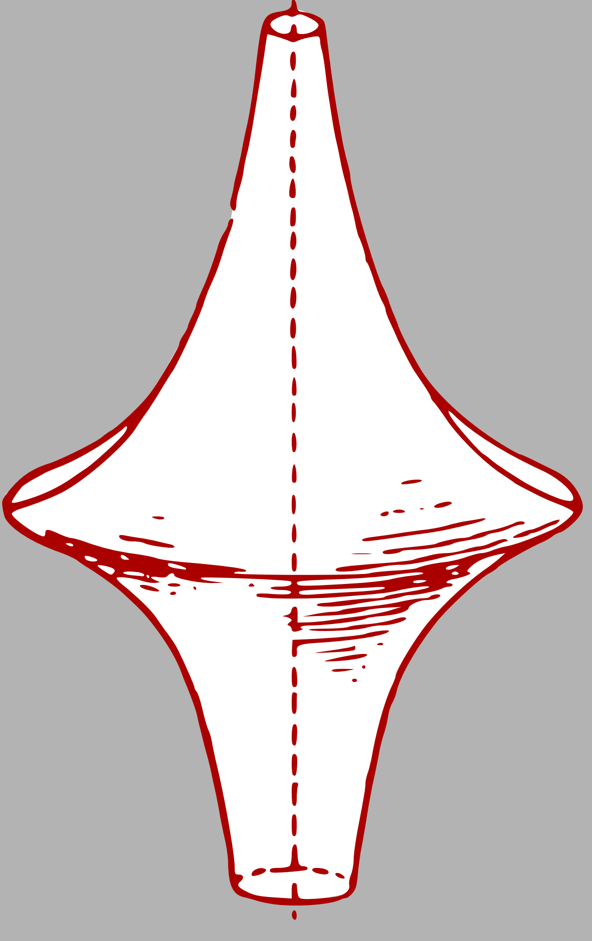
Multiple Integrals Field Theory and Series

B.M. Budak, S.V. Fomin



Mir Publishers Moscow

Multiple Integrals, Field Theory and Series

B.M. Budak, S.V. Fomin

MIR PUBLISHERS MOSCOW

Б. М. Будак, С. В. Фомин

Кратные интегралы и ряды

Издательство «Наука»
Москва

B. M. Budak, S. V. Fomin

Multiple Integrals, Field Theory and Series

An Advanced Course in Higher Mathematics

Translated from the Russian

by

V. M. VOLOSOV, D. Sc.

MIR PUBLISHERS · MOSCOW

UDC 517.3/517.52 (075.8) = 20

First published 1973
Revised from the 1967 Russian edition

На английском языке

$Б \frac{0223-197}{041\,(01)-73}$

Preface

The present book is based on lectures given by the authors at the Physical Department of the Lomonosov State University of Moscow. In the presentation of the material much emphasis has been placed on the physical meaning of mathematical notions and their interrelation as well as on the applications and computational aspects. Chapters 1-6 and Supplement 2 (On Universal Digital Computers) have been written by S.V. Fomin and Chapters 7-11 and Supplement 1 (Asymptotic Expansions) by B.M. Budak. The authors have discussed together the general plan of the book and many details concerning the presentation of the material.

In the preparation of the book the authors have received valuable advice from their colleagues V.A. Ilyin, E.G. Poznyak, A.G. Sveshnikov and others. The authors owe very much to A.N. Tikhonov for his helpful comments and aid. Some important observations have been made by N.V. Yefimov and L.D. Kudryavtsev who have read the manuscript of the book. To all of them the authors express their warmest gratitude.

B.M. Budak

S.V. Fomin

Contents

1 Double Integrals

The definite integral

$$\int_a^b f(x)\,dx$$

is connected with the problems of determining the distance passed over for a given speed, computing the area of a curvilinear trapezoid etc. There are many similar problems involving functions dependent not on one but on many arguments. A typical problem of this kind is to find the volume of a curvilinear cylinder (which is a three-dimensional analogue of a curvilinear trapezoid).

By a *curvilinear cylinder* (*cylindroid*), with base F lying in the x, y-plane, we understand a solid T bounded by the base, a surface $z = f(x, y)$ and the lateral cylindrical surface (Fig. 1.1). It seems

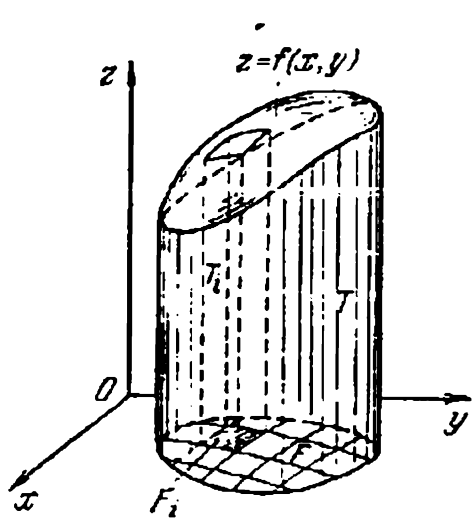


Fig. 1.1

natural to evaluate the volume of such a solid in the following way. Divide the base F by a net of curves into elements, cells, F_i. This results in breaking up the entire cylinder T into elementary cylinders T_i whose bases are the cells F_i. It is clear that the volume of the cylinder T should be understood to be equal to the sum of volumes of the elementary cylinders T_i.

To find the volume of an elementary cylinder T_i, we choose a point (ξ_i, η_i) in F_i and replace the elementary cylinder T_i having a curvi-

linear upper base by a right cylinder with constant altitude $f(\xi_i, \eta_i)$ and the same lower base F_i. In other words, we consider the volume of the elementary cylinder T_i to be approximately equal to

$$f(\xi_i, \eta_i)\,\Delta S_i$$

where ΔS_i is the area of the element F_i. Now we take, as an approximate value of the volume of the whole cylinder T, the sum

$$\sum_{i=1}^{n} f(\xi_i, \eta_i)\,\Delta S_i \tag{1.1}$$

extended over all the cells the base F is divided into. It is intuitively clear that sum (1.1) represents the volume of the cylinder T with degree of accuracy increasing as the sizes of the cells F_i are diminished. To obtain the precise value of the volume we must pass to the limit in expression (1.1) by making the sizes of the elements F_i tend to zero.

This passage to the limit leads to the notion of an integral of a function $f(x, y)$ of two independent variables, i.e. to the *double integral*.

There is an obvious analogy between the above (heuristic) considerations concerning the double integral and the construction of the definite integral of a function of one argument on an interval. The only distinction between them is that in the former we consider functions dependent not on one but on two arguments and that instead of the lengths of subintervals Δx_i we take the areas of the cells F_i into which the figure F, the base of the cylinder, is divided.

Besides the problem of computing the volume of a curvilinear cylinder there are many other problems connected with the concept of the double integral. Some of them will be discussed in § 4 of this chapter.

Some physical and geometrical problems lead to the concept of an integral of a function of three and more independent variables. The next chapter is devoted to these integrals.

The problem of evaluating the volume of a curvilinear cylinder indicates that the notion of a double integral is closely related to the concept of area of a curvilinear plane figure because expression (1.1) involves the areas ΔS_i of curvilinear plane elements F_i into which the base of the cylinder has been broken up. Therefore, although we suppose the reader to be familiar with the concept of area*, we begin this chapter with a brief discussion of the basic properties of area.

---

* E.g. see [8], Chapter 11, § 2.

§ 1. AUXILIARY NOTIONS. AREA OF A PLANE FIGURE

1. Interior and Boundary Points. Domain. We are going to remind the reader of some notions which we shall need in what follows. Let a be a point of the x, y-plane. An *open circle* of radius ε with centre at the point a* is referred to as an **ε-neighbourhood**, or simply a **neighbourhood**, of the point. A point a of a given set A is said to be its **interior** point if a "sufficiently small" ε-neighbourhood of the point a entirely consists of points belonging to the set A. A set whose all points are interior is called an **open set**. An open

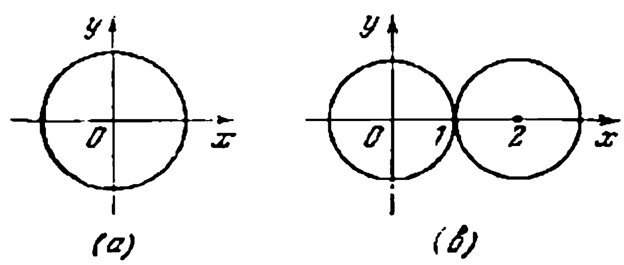


Fig. 1.2

set G is said to be (arcwise) **connected** if each pair of its points can be joined by a continuous curve entirely lying within G. An open connected set is briefly referred to as a **domain**.

For instance, the collection of points whose coordinates satisfy the inequality $x^2 + y^2 < 1$ is a domain (see Fig. 1.2a). The set consisting of the two circles $x^2 + y^2 < 1$ and $(x - 2)^2 + y^2 < 1$ is not a domain because though it is open it is not connected (see Fig. 1.2b).

A point a is called a **boundary point** of the set A if its every neighbourhood contains both points belonging and not belonging to A. A boundary point itself may or may not belong to A. In particular, an open set contains none of its boundary points. The collection of all the boundary points of a set is called its **boundary**. A set containing all its boundary points is called **closed**. Every set can be turned into a closed set (called its **closure**) by adding all the boundary points to it.** In particular, when adding to a domain G all its boundary points we arrive at a set referred to as a **closed domain**.

A point a is called a **limit point** of a set A if in A there exists an infinite sequence of pairwise distinct points $a_1, a_2, \ldots, a_n, \ldots$ convergent to a. A limit point of a set A may or may not be contained in A. A set contains all its boundary points if and only if it is closed. (Prove it.)

---

* That is the totality of all points of the plane whose distances from a are less than ε.

** An arbitrary set may be, of course, neither open nor closed. The collection of all interior points of a set is referred to as its **interior**.—*Tr.*

We say that a set is bounded if it can be placed within a circle of sufficiently large radius. Let A be a bounded set. Denote by $\rho(a_1, a_2)$ the distance between its two arbitrary points. Let now a_1 and a_2 independently run over the whole set A. The set of numbers $\rho(a_1, a_2)$ is then obviously bounded above (because $\rho(a_1, a_2)$ cannot exceed the diameter of the circle in which A is contained). The least upper bound of the set of numbers $\rho(a_1, a_2)$ is called the diameter $d(A)$ of the set A (Fig. 1.3).

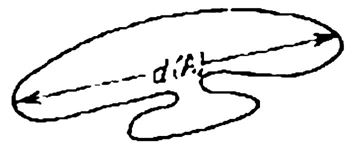


Fig. 1.3

If a set A is a part of a set B or coincides with it (i.e. A is a *subset* of B) we shall designate this fact, as usual, by the symbolic relation $A \subset B$. If a point a belongs to a set A we write $a \in A$.

The union of two sets A and B, i.e. the collection of all points belonging at least to one of the sets, will be denoted as $A + B$, and the **intersection** (or **product**, or **meet**) of two sets A and B which is the collection of all points simultaneously belonging to A and B will be designated by AB.

2. Distance Between Two Sets. Let us introduce another notion which will be applied to the proof of the theorem on the existence of a double integral.

Let A and B be two arbitrary sets in the plane. We shall call the number

$$\rho(A, B) = \inf \rho(a, b) \tag{1.2}$$

the distance between the sets A and B. In (1.2) the greatest lower bound is taken with respect to all the pairs $a \in A$, $b \in B$. We clearly have $\rho(A, B) = 0$ if A and B have at least one point in common. The converse does not hold in the general case; for instance, the distance between the hyperbola $y = \frac{1}{x}$ and the x-axis equals zero although these two lines have no common points at all. At the same time the following theorem (which will be needed in § 2) holds:

Theorem 1.1 (On Separability of Closed Sets). If P and Q are two bounded closed sets with no points in common then $\rho(P, Q) > 0$.

Proof. Assume the contrary, i.e. let $\rho(P, Q) = 0$. Then, by the definition of the distance between two sets, for each $n = 1, 2, \ldots$ there are points $p_n \in P$ and $q_n \in Q$ such that

$$\rho(p_n, q_n) < \frac{1}{n} \tag{1.3}$$

But, $\{p_n\}$ being a bounded infinite sequence, we can choose, according to the well known *Bolzano-Weierstrass theorem* (e.g. see [8], Chapter 14, § 2), a subsequence

$$p_{n_1}, p_{n_2}, \ldots, p_{n_k}, \ldots$$

of $\{p_n\}$ convergent to a point p_0. Then the corresponding points

$$q_{n_1}, q_{n_2}, \ldots, q_{n_k}, \ldots$$

of the sequence $\{q_n\}$ form a subsequence convergent, by (1.3), to the same point p_0.

The point p_0 is sure to belong to the set P. In fact, there are two possibilities here. Either the subsequence $\{p_{n_k}\}$ contains an infinitude of distinct points, and then p_0 is a limit point of P and $p_0 \in P$ because P is closed, or the subsequence $\{p_{n_k}\}$ is stabilized in the

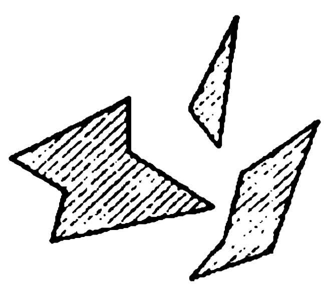


Fig. 1.4

sense that all its points from some number onwards coincide, and then they apparently coincide with p_0 and again $p_0 \in P$. By the same argument, $p_0 \in Q$. But then P and Q have a common point, which contradicts the hypothesis.

Exercise. Show that the theorem remains true when at least one of the two closed sets P and Q is bounded.

3. Area of a Plane Figure. The concept of *area of a polygonal figure* is well known from elementary geometry. (By a *polygonal figure* we mean a set constituted by a finite number of bounded polygons, see Fig. 1.4.) The area of a polygonal figure is a *nonnegative** number possessing the following properties:

1 (*monotonicity*). If P and Q are two polygonal figures and P entirely lies inside Q we have

$$\text{area of } P \leqslant \text{area of } Q$$

2 (*additivity*). If P_1 and P_2 are two polygonal figures without common interior points and $P_1 + P_2$ is the union of the figures

* It can be equal to zero only if the polygonal figure degenerates into a finite number of points or line segments.

we have

$$\text{area of } (P_1 + P_2) = \text{area of } P_1 + \text{area of } P_2{}^*$$

3 (*invariance*). If two polygonal figures P_1 and P_2 are congruent

$$\text{area of } P_1 = \text{area of } P_2$$

Let us now extend the concept of area, preserving the three properties, from polygonal figures to a wider class of plane figures. This problem is solved as follows.

Let F be a *plane figure***. We shall consider all the possible polygonal figures P entirely lying inside F and the polygonal figures Q entirely containing F. The former will be referred to as *embedded figures* and the latter as *enveloping* ones. The areas of embedded figures are bounded above (for instance, by the area of any enveloping figure) and the areas of enveloping figures are bounded below (e.g. by the number zero). Therefore the set of areas of all polygonal figures embedded in the figure F possesses the least upper bound***

$$S_* = S_*(F) = \sup_{P \subset F} \text{ (area of } P)$$

and the set of areas of all enveloping figures possesses the greatest lower bound

$$S^* = S^*(F) = \inf_{Q \supset F} \text{ (area of } Q)$$

The quantity S_* is known as the **interior** (**Jordan******) **content** of the figure F and S^* as its **exterior** (**Jordan**) **content**. The area of any embedded figure not exceeding the area of any enveloping figure, we have

$$S_* \leqslant S^*$$

If $S_* = S^* = S$ their common value S is simply called the **area** (**the Jordan content**) of the figure F. In this case the figure F is said to be **squarable**.

* We can easily verify that the requirements 1 and 2 are not independent because monotonicity of area is implied by its nonnegativity and additivity. Indeed, if a polygonal figure P lies inside a polygonal figure Q we can represent Q as the union of P and a polygonal figure which can be naturally called the difference between the sets Q and P and designated as $Q - P$. Then, by additivity, we have area of Q = area of P + area of $(Q - P)$. But area of $(Q - P) \geqslant 0$ and therefore area of $Q \geqslant$ area of P.

** I.e. a bounded set of points in the plane.

*** If it is impossible to place any polygonal figure within the figure F we put, by definition, $S_* = 0$.

**** Jordan, Camille (1838-1922), a French mathematician.

Thus, we have extended the concept of area from polygons to a wider class of figures.* The retention of the basic properties of area (i.e. additivity, monotonicity and invariance) will be proved in Sec. 4.

We now establish the following necessary and sufficient condition for a figure being squarable which will be of use for our further aims.

Theorem 1.2. A figure F is squarable if and only if for every $\varepsilon > 0$ there exist two polygonal figures $P \subset F$ and $Q \supset F$ such that

$$\text{area of } Q - \text{area of } P < \varepsilon \tag{1.4}$$

Proof. In fact, if such figures exist it follows from

$$\text{area of } P \leqslant S_* \leqslant S^* \leqslant \text{area of } Q$$

that

$$0 \leqslant S^* - S_* < \varepsilon$$

and therefore, since $\varepsilon > 0$ is chosen arbitrarily, we have $S^* = S_*$.

Conversely, if $S^* = S_*$ then, by definition, for any given $\varepsilon > 0$, there is an embedded polygonal figure P and an enveloping figure Q such that

$$S_* - \frac{\varepsilon}{2} < \text{area of } P \leqslant S_*, \qquad S^* \leqslant \text{area of } Q < S^* + \frac{\varepsilon}{2}$$

which implies

$$\text{area of } Q - \text{area of } P < \varepsilon$$

The collection of the points belonging to Q and, simultaneously, not belonging to P is a polygonal figure of area (area of Q — area of P)

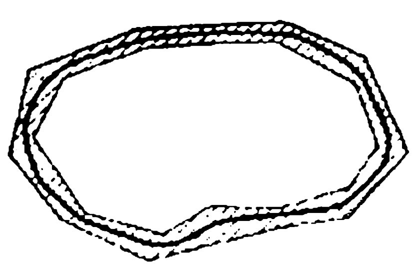


Fig. 1.5

containing the boundary of the figure F. Consequently, the condition of Theorem 1.2 implies that F is squarable if and only if its boundary can be embedded in a polygonal figure of an arbitrarily small area (Fig. 1.5).

---

* Every polygonal figure is obviously squarable, and the new definition of area (introduced with the help of S_* and S^*) yields the original value of its area.

The theorem enables us to establish the squarability of some figures distinct from polygonal ones, for instance, the squarability of the circle. For a circle we can take, as P and Q, a regular inscribed and a regular circumscribed polygon with a sufficiently large number of sides.

By the way, the derivation of the formula for the area of a circle usually performed in elementary courses of geometry is based on the same arguments which are given here in the general form.

Let us introduce the following terminology. We shall say that a set and, in particular, a curve, is of area zero if it can be embedded in a polygonal figure of an arbitrarily small area. This enables us to rephrase Theorem 1.2 as follows:

Theorem 1.2′. For a figure F to be squarable, it is necessary and sufficient that its boundary be of area zero.

Based on the theorem, we now describe a sufficiently wide class of squarable figures to which we shall restrict ourselves in our further considerations.

Lemma. Each rectifiable curve has zero area.*

Proof. Let L be a rectifiable curve of length l. Divide the curve into parts by $n+1$ points so that the length of each part is

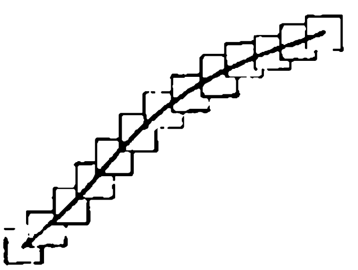


Fig. 1.6

less than $\frac{l}{n}$ (of course, this is always possible) and construct a square of side $\frac{2l}{n}$ with centre at the kth point of division for each $k = 1, 2, \ldots, n+1$ (see Fig. 1.6). The union of the squares is a polygonal figure enveloping the curve L and the area of the polygonal figure does not exceed the sum of the areas of the constituent squares, i.e. is not greater than $\frac{4l^2}{n^2}(n+1)$. Since l is fixed and n can be

* A rectifiable curve is the one that possesses a finite length. As is well known (e.g. see [8], Chapter 11, § 1), if a curve can be represented by parametric equations of the form

$$x = \varphi(t), \quad y = \psi(t), \quad \alpha \leqslant t \leqslant \beta$$

where $\varphi(t)$ and $\psi(t)$ are continuous functions with continuous (or piecewise continuous) derivatives it is rectifiable.

taken as large as desired the curve L can be actually embedded in a figure of an arbitrarily small area. The lemma has been proved.

From the lemma and Theorem 1.2′ we conclude:

Every plane figure (i.e. a bounded plane set) whose boundary is composed of one or several rectifiable curves is squarable.

It is this class of figures that, as a rule, we shall consider in what follows.

Note. We can also point out another class of squarable plane figures. Any curve which can be represented by an equation of the form

$$y = f(x), \quad a \leqslant x \leqslant b$$

where $f(x)$ is a continuous function or by an equation $x = g(y)$, $c \leqslant y \leqslant d$ where $g(y)$ is also continuous is of zero area. (The proof of this fact can be found, for instance, in [8], Chapter 11.) It follows, by Theorem 1.2′, that every figure with a boundary representable as a union of finite number of continuous curves defined by equations of the form $y = f(x)$ or $x = g(y)$ is squarable.

4. Basic Properties of Area. Now we are going to show that the definition of the area of a plane figure thus introduced possesses the properties of monotonicity, additivity and invariance.

Monotonicity is directly implied by the definition of area, and

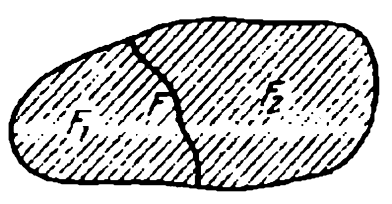


Fig. 1.7

the proof of the property is left to the reader. Let us establish additivity, i.e. prove the following assertion:

(1) *Let F_1 and F_2 be two squarable figures with no interior points in common and F be their union* (see Fig. 1.7); *then F is also squarable and*

$$\text{area of } F = \text{area of } F_1 + \text{area of } F_2 \tag{1.5}$$

The squarability of the figure F follows from Theorem 1.2′ and the fact that the boundary of F is composed of sets of area zero which are some parts of the union of the boundaries of the squarable figures F_1 and F_2.* Therefore, to complete the proof, we must only deduce equality (1.5). To this end consider polygonal figures P_1 and P_2 embedded in F_1 and F_2 and polygonal figures Q_1 and Q_2 enveloping F_1 and F_2, respectively. Since the figures P_1 and P_2

* It appears obvious that every part of a set having zero area is a set of area zero.

do not intersect, the area of the polygonal figure composed of P_1 and P_2 equals area of P_1 + area of P_2. The figures Q_1 and Q_2 (which may intersect) constitute their union Q whose area does not exceed area of Q_1 + area of Q_2. Thus, we have

$$\text{area of } P = \text{area of } P_1 + \text{area of } P_2 \leqslant \text{area of } F \leqslant$$
$$\leqslant \text{area of } Q \leqslant \text{area of } Q_1 + \text{area of } Q_2$$

and

$$\text{area of } P_1 + \text{area of } P_2 \leqslant \text{area of } F_1 + \text{area of } F_2 \leqslant$$
$$\leqslant \text{area of } Q_1 + \text{area of } Q_2$$

Since the differences (area of Q_1 — area of P_1) and (area of Q_2 — — area of P_2) can be made arbitrarily small it follows that equality (1.5) holds. Additivity is thus proved.

Finally, the property of invariance of area immediately follows from the invariance of the area of polygonal figures and from the

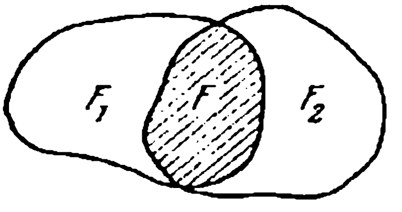


Fig. 1.8

way the area of squarable figures has been introduced by means of the areas of polygonal figures.

There is another property of the squarable figures:

(2) *The intersection of two squarable figures is a squarable figure.*

Indeed, if F_1 and F_2 are two squarable figures and F is their intersection (Fig. 1.8), each boundary point of F is a boundary point at least of one of the figures F_1 and F_2. Therefore our assertion follows from Theorem 1.2′ and the fact that the area of a union of sets of zero area is equal to zero.

5. The Concept of Measure. As has been mentioned above the concept of area has been introduced here according to Jordan's idea. But this way of defining the measure of a set possesses certain disadvantages. Indeed, as has been shown, the union of two squarable figures is squarable. This immediately implies that the union of any finite number of squarable figures is again a squarable figure. But if we take an infinite sequence of squarable figures

$$F_1, F_2, \ldots, F_n, \ldots$$

their union may not be squarable. Here is an example. Take the square

$$0 \leqslant x \leqslant 1, \quad 0 \leqslant y \leqslant 1$$

in the x, y-plane and consider the collection of all its interior points with rational coordinates. It can be easily shown that these points

form a countable set, i.e. they can be arranged into a sequence

$$p_1 = (x_1, y_1), \quad p_2 = (x_2, y_2), \ldots, p_n = (x_n, y_n), \ldots$$

Now take a number $\varepsilon > 0$ and construct a closed circle of radius $r_1 < \frac{\varepsilon}{2}$ with centre at the point p_1 lying within the square. Further, take first of the points $p_2, p_3, \ldots$ which falls outside the circle and construct a new closed circle of radius $r_2 < \frac{\varepsilon}{2^2}$ lying inside the square and not intersecting the former circle. Next we find the first of the remaining points lying outside the circles thus constructed and take it as the centre of a circle of radius $r_3 < \frac{\varepsilon}{2^3}$ contained in the square and not intersecting the circles constructed before. Let us infinitely continue the process in this fashion. We thus obtain an infinite sequence of nonoverlapping circles placed in the square, their union being *everywhere dense* in it.* We can easily show that the union of the circles is a figure F which is nonsquarable in the sense of Jordan (let the reader prove it). On the other hand, it appears natural to attribute to this figure an area equal to the sum of the areas of the circles it is formed of. This sum is obviously equal to

$$\sum_{i=1}^{\infty} \pi r_i^2 < \sum_{i=1}^{\infty} \pi \frac{\varepsilon^2}{2^{2i}} = \frac{1}{3} \pi\varepsilon^2$$

Such difficulties can be avoided by introducing a more flexible and perfect concept of a *Lebesgue measure*** but we cannot discuss it at length here.

§ 2. DEFINITION AND BASIC PROPERTIES OF DOUBLE INTEGRAL

1. Definition of Double Integral. Let us now pass to the main object of this chapter, i.e. the notion of a double integral. Let G be a squarable figure and $f(x, y)$ be a bounded function defined in G. Divide G into a finite number of nonintersecting squarable parts G_i forming a *partition* $\{G_i\}$ of the figure G. Consider a sum of the form

$$\sigma = \sum_{i=1}^{n} f(\xi_i, \eta_i) \Delta S_i \tag{1.6}$$

---

* This means that the union of the circles is a set whose closure coincides with the entire square.

** Lebesgue, Henri Léon (1875-1941), a prominent French mathematician, one of the founders of modern theory of functions.

where ΔS_i is the area of G_i and (ξ_i, η_i) is an arbitrary point belonging to G_i. Sums of form (1.6) will be referred to as **integral sums** (associated with the function $f(x, y)$ and the figure G). We introduce the following definition of the limit of integral sums (1.6).

Definition 1. *Let D be the maximal of the diameters $d(G_i)$ of the figures G_i (the quantity D, the maximal diameter of the partition $\{G_i\}$, is called the* ***fineness of the partition****). A number J is said to be the* ***limit of integral sums*** *(1.6) as $D \to 0$ if for every $\varepsilon > 0$ there is $\delta > 0$ such that*

$$|\sigma - J| < \varepsilon \tag{1.7}$$

when

$$D < \delta \tag{1.8}$$

In other words, inequality (1.7) must hold for all integral sums σ corresponding to the partitions $G = G_1 + G_2 + \ldots + G_n$ which satisfy condition (1.8) irrespective of the way the figure G is broken up into parts G_i and of the particular choice of a point (ξ_i, η_i) in each element of the partition.

Definition 2. *If the limit*

$$\lim_{D \to 0} \sum_{i=1}^{n} f(\xi_i, \eta_i) \Delta S_i$$

of integral sums (1.6) exists it is called the ***double integral of the function*** *$f(x, y)$* ***over the figure*** *G and denoted by the symbol*

$$\iint\limits_G f(x, y)\, ds \quad \text{or} \quad \iint\limits_G f(x, y)\, dx\, dy$$

In this case the function $f(x, y)$, the **integrand**, is said to be **integrable on the figure** G and G is called the **domain of integration**. The expression $f(x, y)\, ds$ or $f(x, y)\, dx\, dy$ is referred to as an **element of integration**.

The notion of a double integral is sometimes introduced in a different manner. A figure G taken from a chosen class of figures is broken into rectangular cells by means of straight lines parallel to the coordinate axes (see Fig. 1.9). In each cell a point (ξ_i, η_i) is then chosen and the sum $\sigma = \sum f(\xi_i, \eta_i) \Delta S_i$ is formed. The sum is taken, say, over all the cells entirely lying within G disregarding those adjoining the boundary of G (the total area of the latter is small). Then the passage to the limit is performed as the maximal diameter of the cells tends to zero. The imperfection of such a defini-

tion is that it is connected with a certain coordinate system in the plane whereas it is intuitively clear that the integral $\iint\limits_G f(x, y)\,ds$, i.e. the volume of the corresponding cylindrical solid, must be independent of the choice of the coordinate system. When the notion of a double integral is introduced by means of such rectangular cells the above fact should be additionally proved but our definition implies it automatically. The definition given here has some other

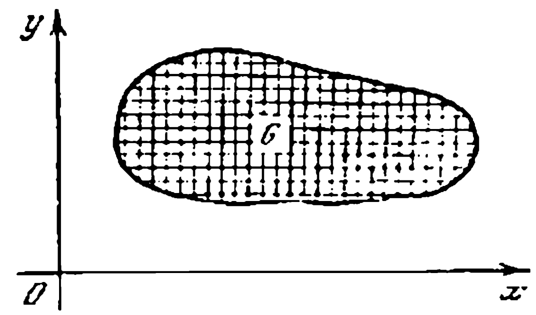


Fig. 1.9

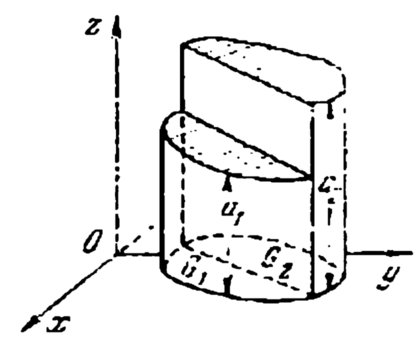


Fig. 1.10

advantages. Let, for instance, a function $f(x, y)$ assume on G only two values: a_1 and a_2 (Fig. 1.10). If the parts G_1 and G_2 on which $f(x, y)$ is equal to a_1 and a_2, respectively, are squarable our definition makes it possible to evaluate the integral $\iint\limits_G f(x, y)\,ds$ without passing to the limit. Intuitively, it is apparent that

$$\iint\limits_G f(x, y)\,ds = (\text{area of } G_1)\cdot a_1 + (\text{area of } G_2)\cdot a_2$$

(prove it). But the definition based on forming rectangular cells would need a sophisticated passage to the limit even in this simple case.

At the same time it should be noted that both definitions result in the same notion of a double integral.

2. Conditions for Existence of Double Integral. Upper and Lower Darboux Sums. Let us find out what requirements should be imposed on a function $f(x, y)$ defined over a squarable figure G in order to guarantee the existence of the double integral

$$\iint\limits_G f(x, y)\,ds$$

In introducing the definition of the double integral we have supposed the corresponding function $f(x, y)$ to be bounded.* At the same time we can easily construct examples indicating that an arbitrary bounded function is by far not always integrable.**

To establish the integrability conditions it is convenient, as in the case of one independent variable, to use the so-called *Darboux**** *upper* and *lower sums*.

Let $f(x, y)$ be a bounded function defined on a squarable figure G, and $\{G_i\}$ be a partition of the figure. Denote by M_i and m_i the least upper and the greatest lower bounds of the values of $f(x, y)$ on the element G_i. The sums

$$\Omega = \sum_{i=1}^{n} M_i \Delta S_i \quad \text{and} \quad \omega = \sum_{i=1}^{n} m_i \Delta S_i$$

are, respectively, referred to as the upper and the lower Darboux sums of the function $f(x, y)$ (corresponding to the given partition $\{G_i\}$ of the figure G). We obviously have $\Omega \geqslant \omega$ for any partition $\{G_i\}$.

Let us enumerate the basic properties of the upper and lower sums.

(1) For every partition $\{G_i\}$ of the figure G, the corresponding upper and lower sums are, respectively, the least upper bound and the greatest lower bound for the integral sums

$$\sum_{i=1}^{n} f(\xi_i, \eta_i) \Delta S_i$$

---

* As is known, a function of one variable which is (Riemann) integrable on an interval is necessarily bounded (e.g. see [8], Chapter 10). But the argument applied to proving this fact cannot be completely extended to the case of two arguments. Actually, when taking different partitions of a squarable figure G into squarable elements G_i we cannot, in general, avoid the cases in which some of the elements are of area zero. But this means that the corresponding integral sums $\sum f(\xi_i, \eta_i) \Delta S_i$ must not necessarily be unbounded for each partition even if the function $f(x, y)$ is not bounded (because the function may turn out to be unbounded only on those elements of partition whose area equals zero). This cannot be the case for a function of one variable when we break up the interval of integration into nonoverlapping half-segments. It is possible to avoid the appearance of elements of area zero for functions of two (or several variables) by restricting both the class of figures and the class of partitions in question. Another way out (which we follow in our presentation of the theory) is to completely exclude unbounded functions.

** An example of a bounded but nonintegrable function of two variables is the one defined on the square $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ in the following way: $f(x, y) = 1$ if x and y are rational numbers and $f(x, y) = 0$ if otherwise. The proof of the fact that the function thus constructed is nonintegrable is left to the reader.

*** Darboux, Jean Gaston, a French mathematician (1842-1917).

associated with the partition $\{G_i\}$ (for all the possible ways of choosing the points (ξ_i, η_i)). In particular, we always have

$$\omega = \sum_{i=1}^{n} m_i \Delta S_i \leqslant \sum_{i=1}^{n} f(\xi_i, \eta_i)\, \Delta S_i \leqslant \sum_{i=1}^{n} M_i \Delta S_i = \Omega$$

Indeed, the inequality

$$\sum_{i=1}^{n} f(\xi_i, \eta_i)\, \Delta S_i \leqslant \sum_{i=1}^{n} M_i \Delta S_i = \Omega$$

obviously holds for any choice of points (ξ_i, η_i) on G_i $(i = 1, 2, \ldots, n)$. On the other hand, by the definition of the least upper bound, for every $\varepsilon > 0$ it is possible to take a point (ξ_i, η_i) in each element G_i of the partition $\{G_i\}$ so that $M_i - f(\xi_i, \eta_i) < \frac{\varepsilon}{S}$ (where S is the area of the domain G). But then we have

$$\Omega - \sum_{i=1}^{n} f(\xi_i, \eta_i)\, \Delta S_i = \sum_{i=1}^{n} (M_i - f(\xi_i, \eta_i))\, \Delta S_i < \frac{\varepsilon}{S} \sum_{i=1}^{n} \Delta S_i = \varepsilon$$

An analogous argument applies in the case of a lower sum.

A partition $\{G'_j\}$ will be referred to as a *refinement* of a partition $\{G_i\}$ if each element G_i of the latter is either an element of the former or a union of several elements of the former partition. In these terms we can formulate the following assertion:

(2) If Ω and ω are the upper and the lower sums corresponding to a partition $\{G_i\}$, and Ω' and ω' are the upper and the lower sums for a refinement $\{G'_j\}$ of $\{G_i\}$, then

$$\omega \leqslant \omega' \leqslant \Omega' \leqslant \Omega$$

that is the upper Darboux sum does not increase and the lower one does not decrease as the partition is refined.

Actually, let $\{G_i\}$ be a partition of the figure G and $\{G'_j\}$ be its refinement. Then each element G_i of the partition $\{G_i\}$ is the union of some elements $G'_{i\alpha}$, $\alpha = 1, 2, \ldots, k_i$ of the latter partition. Furthermore we have

$$\Delta S_i = \sum_{\alpha=1}^{k_i} \Delta S'_{i\alpha} \tag{1.9}$$

$$M_i \geqslant M_{i\alpha}, \qquad \alpha = 1, 2, \ldots, k_i \tag{1.10}$$

each element G'_j being a constituent of only one element G_i. It follows that

$$\Omega = \sum_{i=1}^{n} M_i \Delta S_i \geqslant \sum_{i=1}^{n} \sum_{\alpha=1}^{k_i} M_{i\alpha} \Delta S'_{i\alpha} = \Omega'$$

We similarly prove the inequality $\omega \leqslant \omega'$.

(3) Let $\{G_i'\}$ and $\{G_j''\}$ be two arbitrary partitions of the figure G, and Ω', ω' and Ω'', ω'' be, respectively, the upper and the lower sums associated with the partitions. Then we have

$$\Omega' \geqslant \omega'' \quad \text{and} \quad \Omega'' \geqslant \omega'$$

i.e. every lower sum (corresponding to a given function $f(x, y)$) does not exceed any upper sum (corresponding to the same function). To prove the property we first of all note that for any two partitions of the same figure G there exists their "*common refinement*", i.e. a partition such that it serves as a refinement of each of the two partitions. For instance, to construct such a common refinement we can take, as its elements, the intersections of elements G_i' of one partition with elements G_j'' of the other (of course, we only take those elements G_i' and G_j'' which have common points).

Now consider the upper and the lower sums corresponding to the partitions $\{G_i'\}$, $\{G_j''\}$ and to their common refinement $\{\hat{G}_k\}$. Denote them, respectively, as Ω', ω'; Ω'', ω'' and $\hat{\Omega}$, $\hat{\omega}$. Then, by property (2),

$$\Omega' \geqslant \hat{\Omega}, \quad \Omega'' \geqslant \hat{\Omega}$$

and

$$\omega' \leqslant \hat{\omega}, \quad \omega'' \leqslant \hat{\omega}$$

Besides, we obviously have the inequality

$$\hat{\omega} \leqslant \hat{\Omega}$$

Hence, we have

$$\Omega' \geqslant \hat{\Omega} \geqslant \hat{\omega} \geqslant \omega''$$

and, similarly,

$$\Omega'' \geqslant \hat{\Omega} \geqslant \hat{\omega} \geqslant \omega'$$

The assertion has thus been proved.

The collection of all upper sums corresponding to a given function $f(x, y)$ is bounded below (e.g. an upper sum cannot be less than any lower sum) and the collection of all lower sums is bounded above (e.g. a lower sum cannot exceed any upper sum). Therefore the totality of the upper sums possesses the greatest lower bound which we designate as $\overline{J}$ and the totality of the lower sums has its least upper bound $\underline{J}$. The numbers $\overline{J}$ and $\underline{J}$ are, respectively, called the upper and the lower (Darboux) integrals (corresponding to the domain G and the function $f(x, y)$).

The upper and the lower integrals satisfy the inequality

$$\underline{J} \leqslant \overline{J}$$

In fact, assume the contrary, i.e. $\underline{J} > \overline{J}$. Then there exists a number $\varepsilon > 0$ such that

$$\underline{J} - \overline{J} > \varepsilon > 0 \tag{1.11}$$

Furthermore, by the definition of the least upper and greatest lower bounds, there is an upper sum Ω_1 and a lower sum ω_2 such that

$$\Omega_1 - \overline{J} < \frac{\varepsilon}{2} \quad \text{and} \quad \underline{J} - \omega_2 < \frac{\varepsilon}{2}$$

that is

$$\Omega_1 - \omega_2 + (\underline{J} - \overline{J}) < \varepsilon$$

Consequently, by (1.11), we have

$$\Omega_1 - \omega_2 < 0$$

which contradicts property (3).

Properties (1)-(3) of the upper and the lower sums enable us to establish the following necessary and sufficient condition for the integrability of a function $f(x, y)$ which is completely analogous to the corresponding necessary and sufficient condition for the existence of the definite integral of a function of one argument (e.g. see [8], Chapter 10, Theorem 10.1):

Theorem 1.3. *A bounded function $f(x, y)$ defined on a squarable figure G is integrable over G if and only if for every $\varepsilon > 0$ there exists a partition of the figure G such that the Darboux sums associated with the partition satisfy the condition $\Omega - \omega < \varepsilon$.*

The proof of the theorem is based on the following (Darboux) lemma:

Darboux Lemma. *The upper (lower) integral $\overline{J}$ ($\underline{J}$) is the limit of the upper (lower) Darboux sum as $D \to 0$ (where D is the maximal of the diameters $d(G_i)$ of the elements G_i of the partition $\{G_i\}$ of the figure G).*

For convenience, we introduce the notion of the boundary of a partition. If we are given a partition $\{G_i\}$ of a figure G into squarable parts G_i the union L of the boundaries L_i of all the elements G_i will be referred to as the *boundary of the partition* $\{G_i\}$, i.e.

$$L = L_1 + L_2 + \ldots + L_n$$

The boundaries L_i being of area zero for every partition of the figure G into squarable parts G_i, the boundary L of the partition $\{G_i\}$ has a zero area as well.

The boundary L is the union of a finite number of closed sets L_i and therefore it is also closed. (This is a general property of a union of a finite number of closed sets. Let the reader prove it.)

We now proceed to prove Darboux's lemma.

The proof of Darboux's lemma. By the definition of the upper integral $\bar{J}$, for every $\varepsilon > 0$ there is a partition $\{G_i^*\}$ of the figure G such that the corresponding upper sum Ω^* satisfies the condition

$$0 \leqslant \Omega^* - \bar{J} < \frac{\varepsilon}{2}$$

Embed the boundary L^* of the partition in a polygonal figure Q of area less than $\frac{\varepsilon}{2M}$ where $M = \sup\limits_{(x,\,y)\in G} |f(x, y)|$, so that L^* is strictly contained in it. The boundary L^* and the boundary of the polygonal figure Q are two bounded closed sets having no points in common (see Fig. 1.11). Consequently, by Theorem 1.1, the distance between them is a positive quantity α. Now consider an arbitrary partition $\{G_k\}$ of the figure G for which $D < \alpha$. There is

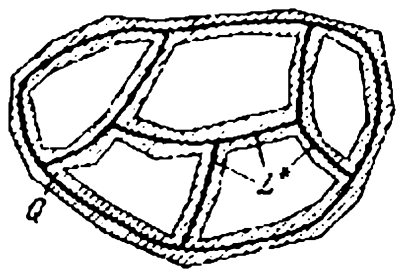


Fig. 1.11

an obvious property of the elements G_k of the partition: if G_k and L^* have at least one common point, then G_k lies entirely in the interior of the figure Q. Such elements G_k will be called *boundary elements*, and all the other will be called *interior elements*. Let us show that, to every partition $\{G_k\}$ with $D < \alpha$, there corresponds an upper sum Ω which differs from $\bar{J}$ by less than ε. To this end, divide the sum Ω into two groups of terms:

$$\Omega = \sum_{k=1}^{n} M_k \Delta S_k = \sum{}' M_k' \Delta S_k' + \sum{}'' M_k'' \Delta S_k''$$

where the summation in $\sum'$ is extended over all interior elements and $\sum''$ is taken over all the boundary elements of the partition $\{G_k\}$. Let us separately estimate each sum. Every interior element of the partition is strictly contained in an element of the partition $\{G_i^*\}$. The corresponding least upper bound M_k' apparently does not exceed the least upper bound of the values of the function $f(x, y)$ assumed on this element of the partition $\{G_i^*\}$. It follows that

$$\sum{}' M_k' \Delta S_k' \leqslant \Omega^*$$

Furthermore, we have the evident inequalities

$$|M_k''| \leqslant M = \sup_{(x,\,y)\in G} |f(x, y)| \quad \text{(for all } k)$$

and

$$\sum'' \Delta S_k'' < \text{area of } Q < \frac{\varepsilon}{2M}$$

Consequently,

$$\left| \sum'' M_k'' \Delta S_k'' \right| < \frac{\varepsilon}{2M}$$

and hence,

$$\Omega = \sum' M_k' \Delta S_k' + \sum'' M_k'' \Delta S_k'' \leqslant \Omega^* + \frac{\varepsilon}{2} < \bar{J} + \varepsilon$$

which is what we set out to prove. The lower sums are considered in a similar way.

Finally, we pass to the proof of Theorem 1.3.

Necessity. Let $f(x, y)$ be integrable and an arbitrary $\varepsilon > 0$ be given. Denote the integral of $f(x, y)$ by the symbol J. From the definition of the limit of integral sums, for any given ε there exists $\delta > 0$ such that for each partition $\{G_i\}$ with $D < \delta$ the inequality

$$\left| J - \sum_{i=1}^{n} f(\xi_i, \eta_i) \Delta S_i \right| < \frac{\varepsilon}{4} \tag{1.12}$$

holds irrespective of the choice of the points (ξ_i, η_i). We also know that the upper and the lower sums corresponding to the partition $\{G_i\}$ are the least upper and the greatest lower bounds of the integral sums associated with the partition. Therefore, we can take a fixed partition and choose the points (ξ_i', η_i') and (ξ_i'', η_i'') within the elements G_i of the partition so that the following inequalities are fulfilled:

$$\Omega - \sum_{i=1}^{n} f(\xi_i', \eta_i') \Delta S_i < \frac{\varepsilon}{4}; \quad \sum_{i=1}^{n} f(\xi_i'', \eta_i'') \Delta S_i - \omega < \frac{\varepsilon}{4} \tag{1.13}$$

Each of the two integral sums satisfying condition (1.12), we deduce, from (1.13), the desired result:

$$\Omega - \omega < \varepsilon$$

Sufficiency. If for every $\varepsilon > 0$ there exists a partition such that

$$\Omega - \omega < \varepsilon$$

we obviously have

$$\bar{J} = \underline{J}$$

Denote the common value of the quantities $\bar{J}$ and $\underline{J}$ by J. Let us show that J is the limit of integral sums, i.e. the double integral of the function $f(x, y)$ over the domain G. By Darboux's lemma

J is the common limit of the upper and lower sums for $D \to 0$. But since the value of any integral sum associated with a partition is contained between the corresponding Darboux sums Ω and ω the number J is the limit of the integral sums as $D \to 0$. The theorem has been proved.

3. Some Important Classes of Integrable Functions. Applying Theorem 1.3, we shall now establish the integrability of some important classes of functions, and, first of all, continuous functions. In what follows we shall regard each function in question as being defined on a bounded closed squarable domain.

Theorem 1.4. Every continuous function $f(x, y)$ defined in a bounded closed domain G is integrable on G.*

Proof. Since $f(x, y)$ is continuous on a bounded closed set it is bounded and uniformly continuous on it.** The uniform continuity of the function $f(x, y)$ implies that for every $\varepsilon > 0$ there is $\delta > 0$ such that if the figure G is divided into parts G_i whose diameters are less than δ the *oscillation of the function* $f(x, y)$ on each of the parts, i.e. the difference $M_i - m_i$, is less than ε. But then

$$\Omega - \omega = \sum_{i=1}^{n} M_i \Delta S_i - \sum_{i=1}^{n} m_i \Delta S_i < \varepsilon \sum_{i=1}^{n} \Delta S_i = \varepsilon S$$

and hence the function $f(x, y)$ is integrable.

The condition of continuity of the integrand is too restrictive. Therefore the theorem below guaranteeing the existence of the double integral for a class of discontinuous functions is important for applications.

Theorem 1.4′. If a function $f(x, y)$ is bounded over a bounded closed domain G and is continuous throughout G possibly except a set of area zero the function is integrable on G.

Proof. Take an arbitrary $\varepsilon > 0$. By the hypothesis, $f(x, y)$ is bounded, that is there exists a number K such that $|f(x, y)| < K$. Let us embed the set on which the function $f(x, y)$ is discontinuous in a polygonal figure Q of area less than $\frac{\varepsilon}{4K}$ (see Fig. 1.12) so that it should be strictly contained within the figure. Denote as $\widetilde{G}$ the part of the domain G not entering into the interior of Q. The boundary points of the polygonal figure Q which belong to G lie in $\widetilde{G}$,

---

* And, of course, squarable. In what follows we shall suppose, without any further stipulation, that the condition of squarability is always fulfilled.

** E.g. see [8], Chapter 14, Theorems 14.6 and 14.8.

and therefore $\tilde{G}$ is closed. The function $f(x, y)$ is continuous on the closed set $\tilde{G}$ and hence is uniformly continuous on it. Choose $\delta > 0$ so that the oscillation of the function $f(x, y)$ on any part of the figure $\tilde{G}$ with diameter less than δ should be less than $\frac{\varepsilon}{2S}$ (where S is the area of G). Now, consider a partition $\{G_i\}$ of the domain G whose first element G_1 coincides with Q and all the other elements

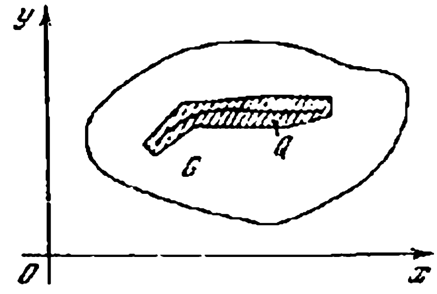


Fig. 1.12

are of diameters less than δ. Let us estimate the difference $\Omega - \omega$ for this partition. We have

$$\Omega - \omega = M_1 \Delta S_1 - m_1 \Delta S_1 + \sum_{i=2}^{n} (M_i - m_i) \Delta S_i <$$

$$< (M_1 - m_1) \frac{\varepsilon}{4K} + \sum_{i=2}^{n} \frac{\varepsilon}{2S} \Delta S_i$$

But $M_1 - m_1 \leqslant 2K$ and $\sum\limits_{i=2}^{n} \Delta S_i < S$, and thus

$$\Omega - \omega < 2K \frac{\varepsilon}{4K} + \frac{\varepsilon}{2S} S = \varepsilon$$

The number $\varepsilon > 0$ being chosen arbitrarily, the function $f(x, y)$ is integrable by virtue of Theorem 1.3.

4. Properties of Double Integral. The basic properties of the double integral are completely analogous to the corresponding properties of the definite integral of a function of one independent variable and therefore we shall only enumerate them without giving the proofs.

1. If functions $f_1(x, y)$ and $f_2(x, y)$ are integrable over a domain G their sum (difference) is also integrable on G and

$$\iint_G [f_1(x, y) \pm f_2(x, y)]\, ds = \iint_G f_1(x, y)\, ds \pm \iint_G f_2(x, y)\, ds$$

2. If k is a constant number and a function $f(x, y)$ is integrable on G the function $kf(x, y)$ is also integrable on G and

$$\iint\limits_G kf(x, y)\,ds = k\iint\limits_G f(x, y)\,ds$$

These two properties express *linearity* of the integral.

3. If a domain G is a union of two domains G_1 and G_2 and a function $f(x, y)$ is integrable on G_1 and G_2 then the function is as well integrable on G. If, besides, G_1 and G_2 have no interior points in common we have

$$\iint\limits_G f(x, y)\,ds = \iint\limits_{G_1} f(x, y)\,ds + \iint\limits_{G_2} f(x, y)\,ds$$

This property is referred to as *additivity* of the integral.

4. If $f_1(x, y)$ and $f_2(x, y)$ are integrable on G and $f_1(x, y) \leqslant \leqslant f_2(x, y)$ then

$$\iint\limits_G f_1(x, y)\,ds \leqslant \iint\limits_G f_2(x, y)\,ds$$

This property is called *monotonicity* of the integral; it implies properties 5 and 6.

5 (*estimation of the modulus of the integral*). If $f(x, y)$ is integrable on G the function $|f(x, y)|$ is also integrable on G and

$$\left|\iint\limits_G f(x, y)\,ds\right| \leqslant \iint\limits_G |f(x, y)|\,ds$$

6 (*mean value theorem*). If a function $f(x, y)$ is integrable on G and satisfies the inequalities

$$m \leqslant f(x, y) \leqslant M$$

we have

$$mS \leqslant \iint\limits_G f(x, y)\,ds \leqslant MS \tag{1.14}$$

where S is the area of the figure G.

The assertion immediately follows from property 4 and an obvious relation

$$\iint\limits_G c\,ds = cS, \qquad c = \text{const}$$

If $f(x, y)$ is a continuous function the mean value theorem can be stated as follows:

6′. In the domain G, there is a point (ξ, η) such that

$$\iint\limits_G f(x, y)\,ds = f(\xi, \eta)\,S \tag{1.15}$$

Indeed, take respectively, as M and m, the least upper bound and the greatest lower bound of the values of the function $f(x, y)$ on G. Then, according to (1.14), we have

$$m \leqslant \frac{1}{S} \iint_G f(x, y)\, ds \leqslant M$$

But, as is well known, a continuous function defined in a closed domain assumes, at some points of the domain, the values equal to its least upper bound M and greatest lower bound m (e.g. see [8], Chapter 14, § 3). Suppose, for simplicity, that the function $f(x, y)$ takes on the values M and m at the points (x_1, y_1) and (x_2, y_2) lying in the interior of the domain G (the argument becomes a little more sophisticated if one of the points or both fall in the boundary of the domain G). Every two points of a domain can be joined by a broken line contained in the domain. Let us connect, by a broken line contained in the domain G, the points (x_1, y_1) and (x_2, y_2) at which the function is, respectively, equal to M and m. The function $f(x, y)$ is continuous along this polygonal line and, consequently, together with the values M and m, it assumes all the intermediate values. In particular, we can find a point (denote it by (ξ, η)) at which

$$f(\xi, \eta) = \frac{1}{S} \iint_G f(x, y)\, ds$$

and thus formula (1.15) has been proved.

§ 3. ADDITIVE SET FUNCTIONS. DERIVATIVE OF A SET FUNCTION WITH RESPECT TO AREA

1. Point Functions and Set Functions. The notion of a function is one of the most important in analysis. We have already dealt with functions dependent on one, two or several arguments. Applying geometrical terminology we can say that such functions are variable quantities dependent on a point of the line (for one argument), on a point in the plane (for two arguments), on a point of a three-dimensional space (for three arguments) or on a point belonging to a space of higher dimension. But in mathematical analysis and its physical applications we often encounter functions of different type for which the values of their arguments are not separate points but certain sets, for instance, some plane or space geometric figures. Functions of this type are known as **set functions**.

As an example of a set function, we can take the area $S(G)$ of a domain* G defined, in a manner described in § 1, for all squarable

* The term "domain" is understood here in a wider sense, not as an open connected set but as a synonym for the term "set". A class of sets on which set functions are considered can be chosen in an arbitrary fashion. In this book, as a rule, we deal with set functions whose argument is a squarable plane figure (domain).

domains in the plane. Take another example. Let a mass be distributed in the x, y-plane. Then, to each domain G lying in the plane, there corresponds a certain number, i.e. a mass μ (G) concentrated on G. Here again we have a variable quantity dependent on a domain, that is a set function.

Now we introduce an important definition.

Definition. *A set function F (G) is said to be* ***additive*** *if the following conditions hold:*

(1) if F (G) is defined for domains G_1 and G_2 it is as well defined for their union $G_1 + G_2$;

(2) if G_1 and G_2 have no interior points in common we have

$$F(G_1 + G_2) = F(G_1) + F(G_2)^*$$

The above functions, area and mass, possess these properties. We can give many other examples of additive set functions: surface charge, amount of light energy impinging on an illuminated surface, fluid pressure acting upon the bottom of a vessel etc.

We can also indicate examples of nonadditive set functions. For instance, if, with every squarable domain, we associate the square of its area we obtain a set function which is not additive.

Additive functions whose argument is not a plane but a space figure will be treated in the next chapter devoted to the triple integral.

2. Double Integral as an Additive Function of Its Domain of Integration. Let us consider the double integral

$$\iint\limits_G f(x, y)\, ds$$

in which the integrand f (x, y) is regarded as being fixed whereas the domain of integration G is variable. Then the integral becomes a function Φ (G) of the domain G. By virtue of property 2 of the double integral (see the foregoing section), this function is additive. As a class of sets for which the function is defined, we can take the totality of all squarable figures contained in the domain G_0 on which f (x, y) is defined.

3. Derivative of a Set Function with Respect to Area. Take again a function of type μ (G), i.e. a mass distribution in the plane. If G is a squarable domain and S (G) its area the ratio

$$\frac{\mu(G)}{S(G)} \tag{1.16}$$

* In particular, it follows that if G is of area zero, then F $(G) = 0$. In the case of a mass this means that we only consider mass distributions having a two-dimensional (surface) density (but not concentrated at separate points or curves).

is the *mean density* of mass distribution in the domain G. We now infinitely diminish the sizes of the domain G by contracting it to a fixed point p_0. If, in this process, ratio (1.16) tends to a limit $\rho(p_0)$ the limit is called the *density of mass distribution at the point* p_0. Thus, a mass distribution in the plane can be directly defined by indicating an additive set function $\mu(G)$ or characterized by the corresponding density which is a point function.

We now pass from our concrete example (mass distribution) to an arbitrary set function. Unlike mass distribution, an arbitrary set function can assume both positive and negative values.

Let $F(G)$ be an additive set function defined for all the squarable domains*. We say that a number A is the *limit* of the ratio

$$\frac{F(G)}{S(G)}$$

(where $S(G)$ is the area of the domain G), as the domain G is contracted to a point p_0, if for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$\left|\frac{F(G)}{S(G)} - A\right| < \varepsilon$$

for each domain G entirely lying in the circle of radius δ with centre at the point p_0.

This limit will be denoted by the symbol

$$\lim_{G \to p_0} \frac{F(G)}{S(G)} \text{ or } \frac{dF}{ds}$$

and referred to as the derivative of the additive set function $F(G)$ with respect to area. The derivative is not a set function but an ordinary point function, i.e. a variable quantity dependent on a point.

Turning back to the above example, we can say that the density $\rho(p_0)$ of a mass distribution in the plane is the derivative of mass with respect to area.

4. Derivative of a Double Integral with Respect to the Area of Its Domain of Integration. The mean value theorem for the double integral (see § 2, Sec. 4, property 6) implies the following result.

Take the integral

$$F(G) = \iint_G f(x, y)\, ds \tag{1.17}$$

where $f(x, y)$ is a fixed function which is supposed to be continuous throughout a chosen part of the plane. Let us show that the additive

* Or for all squarable domains contained in a fixed domain.

set function defined by relation (1.17) possesses the derivative with respect to area which coincides with the integrand $f(x, y)$.

Actually, let p_0 be a fixed point and G be a domain lying within a circle with centre at p_0. Denote by m and M the greatest lower bound and the least upper bound of the values of the function $f(x, y)$ in the domain G. By virtue of the mean value theorem, we have

$$m \leqslant \frac{1}{S(G)} \iint_G f(x, y)\, ds \leqslant M$$

When the domain G is contracted to the point p_0, i.e. when the radius of the circle tends to zero, the numbers m and M tend, because of the continuity of $f(x, y)$ at the point p_0, to the same value, namely, to the value taken by the function $f(x, y)$ at the point. Consequently, the ratio whose values lie between m and M tends to the same limit. Hence, we really have

$$\frac{dF}{ds} = f(x, y)$$

5. Reconstruction of an Additive Set Function from Its Derivative. We have discussed the problem of finding the derivative of a set function. Here we shall consider the reverse problem: let a point function $f(x, y)$ be given and let it be necessary to determine a set function $F(G)$ whose derivative coincides with $f(x, y)$. If the function $f(x, y)$ is continuous we can immediately indicate such a set function, namely, the double integral

$$\iint_G f(x, y)\, ds \tag{1.18}$$

regarded as a function of G. It appears natural to pose the question on whether there exist some other set functions with the same derivative. Let us show that if $f(x, y)$ is continuous there is only one additive set function whose derivative is $f(x, y)$ (and which thus is expressible in the form of double integral (1.18)).

If $F_1(G)$ and $F_2(G)$ are two additive set functions with the same derivative with respect to area we have

$$\frac{d}{ds}(F_1 - F_2) \equiv 0$$

It is therefore sufficient to prove the following assertion:

If $\frac{dF}{ds} \equiv 0$, then $F \equiv 0$. The proof is implied by the lemma stated below.

Lemma. If the derivative $\frac{dF}{ds}$ of an additive set function $F(D)$ exists in a bounded closed domain D and is nonnegative, then $F(D) \geqslant 0$.

Proof. Assume the contrary, i.e. let $F(D) < 0$. Then there is $l < 0$ such that

$$\frac{F(D)}{S(D)} \leqslant l < 0$$

that is

$$F(D) \leqslant lS(D) \tag{1.19}$$

Further, take a sequence of positive numbers $\varepsilon_1, \varepsilon_2, \ldots$ convergent to zero and break up the domain D into a finite number of parts D_i so that their diameters are less than ε_1. Then at least for one of these parts (denote it as $D^{(1)}$) we must have

$$F(D^{(1)}) \leqslant lS(D^{(1)})$$

because if the opposite inequality

$$F(D_i) > lS(D_i)$$

were fulfilled for all D_i we should sum up these inequalities over all D_i and thus arrive at a contradiction to inequality (1.19).

Now, divide $D^{(1)}$ into parts with diameters less than ε_2. Among them there is at least one (denote it by $D^{(2)}$) for which the inequality

$$F(D^{(2)}) \leqslant lS(D^{(2)})$$

holds. Continuing in this manner we obtain the sequence $\{\overline{D}^{(n)}\}$ which is a *nested collection* of closed and bounded domains (the symbol $\overline{D}^{(n)}$ designates the closure of $D^{(n)}$, and we apparently have $F(\overline{D}^{(n)}) = F(D^{(n)})$). The diameters of $\overline{D}^{(n)}$ tending to zero, there exists a single point belonging to all $\overline{D}^{(n)}$ * (let us denote this point by p_0). By the hypothesis, the derivative $\frac{dF}{ds}$ exists everywhere in D, and, in particular, at the point p_0, and therefore its value at the point can be expressed as

$$\lim_{n\to\infty} \frac{F(D^{(n)})}{S(D^{(n)})} \tag{1.20}$$

But, according to the construction of the sequence $\{D^{(n)}\}$, the ratio $\frac{F(D^{(n)})}{S(D^{(n)})}$ does not exceed the fixed negative number l for all n, and thus limit (1.20) cannot be nonnegative. The lemma has been proved.

Replacing $F(G)$ by $-F(G)$ and applying the lemma we see that if $\frac{dF}{ds}$ exists and is nonpositive, we have $F(D) \leqslant 0$. Finally, if

$$\frac{dF}{ds} = 0$$

* This is the two-dimensional analogue of the *nested interval theorem* (e.g. see [8], Chapter 3, § 3).

that is if we simultaneously have

$$\frac{dF}{ds} \geqslant 0 \quad \text{and} \quad \frac{dF}{ds} \leqslant 0$$

then $F(D) = 0$ for every bounded closed domain.

6. Definite Integral of a Function of One Argument as a Function of Its Interval of Integration. Let us now compare what has been said with the analogous facts of the theory of the definite integral of a function of one independent variable. The definite integral

$$\int_a^b f(\xi)\, d\xi$$

can be regarded (for a fixed function f) as a function of the interval $[a, b]$, i.e. as a set function on the line. Furthermore, the well known properties of the definite integral imply the additivity of this set function. But a line segment is completely specified by two points, namely, by its end points. If one of the end points is fixed a function of the line segment becomes an ordinary point function. We encounter this particular situation when we consider the integral

$$\int_a^x f(\xi)\, d\xi \tag{1.21}$$

(for a fixed a) as a function of its upper limit of integration. If we substitute another point a' for the lower limit a function (1.21) gains a constant (independent of x) increment, namely, the one equal to

$$\int_a^{a'} f(\xi)\, d\xi$$

Thus, an integral of a function of one argument is a uniquely specified set function on the line. When we regard such a function as a function of intervals it can be reduced to a function of one independent variable determined to within an arbitrary (additive) constant. The theorem on the derivative of a double integral with respect to area and the theorem on the reconstruction of a set function from its derivative are, respectively, the two-dimensional analogues of the theorem on the derivative of a definite integral of a continuous function with respect to its upper limit and of the one asserting that an antiderivative is determined to within an arbitrary constant summand.

7. Extension of Additive Set Functions. If a function is not given for all possible values of its argument for which it may be defined we can usually extend the function if some of its properties are

known. For example, if a function $f(x)$ is known to be linear, i.e. is of the form

$$f(x) = ax + b$$

then, to find its values everywhere, it is sufficient to have at one's disposal the values of the function at any two distinct points. If $f(x)$ is a periodic function of a period $T > 0$, i.e. if it possesses the property

$$f(x + T) \equiv f(x)$$

for all x, then, to find its values everywhere, it is sufficient to know the values of the function on the interval $[0, T]$. For instance, if the values of $\sin x$ are known for all x from 0 to 2π we can find the sine of any angle. The set functions can be treated analogously. If a set function $F(G)$ is known to be additive and if its values assumed on a certain class of sets are given we can sometimes uniquely extend the function, preserving its additivity, to a wider class of sets. For example, let $F(G)$ be an additive set function defined on all triangles. Then it can be extended, as an additive function, to all polygons (and then to a wider class of sets).

We have dealt with a problem of this type in § 1 where we have studied the notion of area. The area is an additive set function of a domain which has been originally regarded as being defined on polygons (or polygonal figures) and then extended, with preservation of its additivity, to a wider class of figures which we have called squarable.

The general problem of constructing an additive extension of a set function and determining the widest class of sets for which the function can be defined plays an important role in many divisions of mathematics. But here we shall not discuss these questions in detail because it would involve the introduction and systematic application of ideas and concepts of the general theory of measure.

§ 4. SOME PHYSICAL AND GEOMETRICAL APPLICATIONS OF THE DOUBLE INTEGRAL

1. Evaluating Volumes. At the beginning of this chapter we have already discussed a geometrical problem leading to the notion of a double integral, that is the problem of finding the volume of a curvilinear cylinder. We have seen that, for a cylindroid bounded below by a closed domain G in x, y-plane and above by a surface $z = f(x, y)$ where $f(x, y)$ is a nonnegative continuous function, the integral sum

$$\sum_{i=1}^{n} f(\xi_i, \eta_i)\,\Delta S_i \tag{1.22}$$

gives an approximate value of the volume. (The sum is taken over all elements G_i of a partition of the figure G into squarable parts,

ΔS_i is the area of the element G_i, and $(\xi_i, \eta_i) \in G_i$.) As has been said in the introduction to this chapter, the exact value of the volume equals the limit to which integral sums (1.22) tend as the fineness of the partition tends to zero. But the limit of sums (1.22) is nothing but the double integral of the function $f(x, y)$ over G. Its existence (under certain assumptions concerning $f(x, y)$ and G) has already been proved (Theorem 1.3). Hence, the volume V of a curvilinear cylinder bounded below by a closed domain G and above by a surface $z = f(x, y)$ (where $f > 0$ is continuous) is represented by the double integral

$$\iint_G f(x, y)\, ds$$

Strictly speaking, the volume of a curvilinear cylinder must be defined as the value of the double integral. The concept of the volume, clear though it may be from the geometrical point of view, is not given beforehand and therefore our considerations only indicate that such a definition looks natural and is coherent with geometric intuition.

We shall consider here some other problems to which the notion of double integral is applied.

2. Computing Areas. Assuming that the integrand $f(x, y)$ of a double integral is identically equal to unity we arrive at the expression

$$\iint_G ds \tag{1.23}$$

which is obviously equal to the area of the figure G because each integral sum corresponding to integral (1.23) equals that area. The formula

$$S = \iint_G ds \tag{1.24}$$

for computing the area is sometimes more convenient than the well known formula

$$S = \int_a^b f(x)\, dx$$

expressing the area of a curvilinear trapezoid because formula (1.24) is applicable not only to a curvilinear trapezoid but also to any squarable figure occupying an arbitrary position with respect to the coordinate axes.

3. Mass of a Plate. Consider a plate lying in the x, y-plane, i.e. a domain G in which a mass with surface density $\rho(x, y)$ is distribut-

ed. Let us find the mass of the plate from the given density $\rho(x, y)$ under the hypothesis that $\rho(x, y)$ is a continuous function in x and y. Break up G into parts G_i in an arbitrary way and take a point (ξ_i, η_i) in each of the parts. The mass of each element G_i can be approximately regarded as equal to $\rho(\xi_i, \eta_i)\, \Delta S_i$ (where ΔS_i is the area of G_i) and the total mass of the plate as equal to the sum

$$\sum_{i=1}^{n} \rho(\xi_i, \eta_i)\, \Delta S_i \tag{1.25}$$

taken over all the elements of the partition. To obtain the exact value of the mass of the plate it is necessary to pass to the limit in the sum as the maximal diameter of the partition $\{G_i\}$ of the domain G is infinitely diminished. Then expression (1.25) turns into the double integral

$$\iint_G \rho(x, y)\, ds \tag{1.26}$$

which gives the mass of the plate.

It appears clear that determining the mass of a plate from its density is a particular case of the general problem of reconstructing a set function from its derivative which has been discussed above (see § 3).

4. Coordinates of the Centre of Gravity of a Plate. Let us determine the coordinates of the centre of gravity of a plate occupying a domain G in the x, y-plane. Suppose that $\rho(x, y)$ is the density of the plate at the point (x, y). Divide the domain G into parts G_i, choose a point (ξ_i, η_i) in each of the parts and consider the mass of each part to be approximately equal to $\rho(\xi_i, \eta_i)\, \Delta S_i$ where ΔS_i is the area of the subdomain G_i. Each mass can be thought of as being concentrated at one point, namely, at the point (ξ_i, η_i). Then we can write the well known expressions for the coordinates x_c and y_c of the centre of gravity of the system of material points:

$$x_c = \frac{\sum\limits_{i=1}^{n} \xi_i \rho(\xi_i, \eta_i)\, \Delta S_i}{\sum\limits_{i=1}^{n} \rho(\xi_i, \eta_i)\, \Delta S_i}; \qquad y_c = \frac{\sum\limits_{i=1}^{n} \eta_i \rho(\xi_i, \eta_i)\, \Delta S_i}{\sum\limits_{i=1}^{n} \rho(\xi_i, \eta_i)\, \Delta S_i} \tag{1.27}$$

Expressions (1.27) are approximately equal to the coordinates of the centre of gravity of the plate. To receive the exact values of the coordinates we must pass to the limit in formulas (1.27) as the partition is infinitely refined, i.e. $D \to 0$. Then the sums entering into formulas (1.27) turn into the corresponding integrals and thus we obtain the formulas for the coordinates of the centre of gravity

of the plate:

$$x_c = \frac{\iint\limits_G x\rho(x, y)\,ds}{\iint\limits_G \rho(x, y)\,ds}; \qquad y_c = \frac{\iint\limits_G y\rho(x, y)\,ds}{\iint\limits_G \rho(x, y)\,ds} \tag{1.28}$$

If the plate is homogeneous, i.e. $\rho = \text{const}$, the formulas for the coordinates of the centre of gravity are simplified:

$$x_c = \frac{\iint\limits_G x\,ds}{\iint\limits_G ds}; \qquad y_c = \frac{\iint\limits_G y\,ds}{\iint\limits_G ds} \tag{1.29}$$

5. Moments of Inertia of a Plate. As is well known, the moment of inertia of a material point about an axis is equal to the product of the mass of the point by the square of its distance from the axis and the moment of inertia of a system of material points (about the same axis) equals the sum of the moments of inertia of the mass points it is formed of. Let a domain G in the x, y-plane be occupied by a plate of density $\rho(x, y)$. Break up the domain G into parts G_i with areas ΔS_i, choose a point (ξ_i, η_i) in every part and replace the plate by the system of masses $\rho(\xi_i, \eta_i)\,\Delta S_i$ concentrated at the points (ξ_i, η_i). Then the moment of inertia of this system of material points about the y-axis is equal to

$$\sum_{i=1}^{n} \xi_i^2 \rho(\xi_i, \eta_i)\,\Delta S_i$$

This expression is taken as an approximate value of the moment of inertia of the plate, and the smaller the diameter of the partition, the greater the accuracy of the approximation. Passing then to the limit as the maximal diameter of the partition of the domain G tends to zero we receive the following formula for the moment of inertia of the plate about the y-axis:

$$I_y = \iint\limits_G x^2 \rho(x, y)\,ds \tag{1.30}$$

Similarly, the moment of inertia of the plate about the x-axis is equal to

$$I_x = \iint\limits_G y^2 \rho(x, y)\,ds \tag{1.31}$$

Now let us find the moment of inertia I_0 of the plate about the origin of the coordinate system. Taking into account that the moment of inertia of a material point of mass m (placed at the point (x, y))

about the origin is equal to

$$m\,(x^2 + y^2)$$

and applying the same arguments we find that

$$I_0 = \iint\limits_G (x^2 + y^2)\,\rho(x, y)\,ds$$

i.e.

$$I_0 = I_x + I_y$$

6. Luminous Flux Incident on a Plate. Let a plate occupying a domain G of the x, y-plane be illuminated by a point source of light placed at a point with the coordinates $(0, 0, z_0)$. Suppose that the light intensity of the source is the same in all directions and denote it by I. Let us compute the luminous flux incident on the plate.

The luminous flux dF impinging on an elementary area ds is equal to $I\,d\omega$ where $d\omega$ is the solid angle at the point $(0, 0, z_0)$ subtended by the surface ds. Furthermore, $d\omega$ is equal to the product of the ratio of the area ds to the square of its distance from the source by the cosine of the angle between the normal to the area and the direction from the area to the source. The value of the derivative $\frac{dF}{ds}$ at a point (x, y) of the plate is known as the *intensity of illumination* at the point (denote it by $A\,(x, y)$). It follows that

$$A\,(x, y) = \frac{dF}{ds} = \frac{I\,d\omega}{ds} = \frac{Iz_0}{(x^2 + y^2 + z_0^2)^{3/2}}$$

The total luminous flux falling on the plate is equal to the double integral of $A\,(x, y)$ over the domain G, i.e. equal to

$$Iz_0 \iint\limits_G \frac{ds}{(x^2 + y^2 + z_0^2)^{3/2}}$$

7. Flux of a Fluid Through the Cross Section of a Channel. Consider a fluid flow in a channel, and take a cross section of the channel perpendicular to the direction of the flow. Introducing a Cartesian coordinate system x, y in the plane of the section we can regard the speed V of the fluid, at each point of the section, as a function of x and y, i.e. $V = V\,(x, y)$. Let us compute the amount of the fluid passing across the section in unit time. Take an infinitesimal element ds of the section. The quantity of fluid passing through the element in unit time is obviously equal to the mass of the elementary fluid cylinder with base ds and altitude $V\,(x, y)$, that is equal to

$$\rho V\,(x, y)\,ds \tag{1.32}$$

where ρ is the density of the fluid. To find the amount of fluid passing through the whole section we must sum up infinitesimal elements

(1.32), i.e. take the double integral

$$\iint\limits_G \rho V(x, y)\, ds$$

over the section.

Note. In above considerations and, particularly, in the last problem we have used such terms as "an infinitesimal element of area", "an element of mass" and the like. This terminology is widely applied, especially in physical literature. For instance, we say that, for a plate with density $\rho(x, y)$, the quantity

$$\rho(x, y)\, ds$$

is its "element of mass" (concentrated on "an element of area ds") and the total mass of the plate, that is the integral

$$\iint\limits_G \rho(x, y)\, ds$$

is regarded as "the sum of the mass elements".

The meaning of such statements is that we always imply the corresponding processes of passing to the limit (from finite sums to integrals) which have been encountered in the above problems. In what follows we shall sometimes use this "physical" language (keeping in mind its real sense based on the corresponding passage to the limit).

§ 5. REDUCING DOUBLE INTEGRAL TO A TWOFOLD ITERATED INTEGRAL

We have already discussed the definition and basic properties of the double integral, the conditions for its existence and some physical and geometrical problems involving this notion but we have not yet studied the practical ways of evaluating double integrals. The most important role in the solution of this problem is played by the theorem asserting that the evaluation of a double integral can be reduced, under some general suppositions, to successive separate integrations with respect to each variable. It is the proof of the theorem that we are going to study in § 5.

1. Heuristic Considerations. The basic idea of the theorems proved below lies in the following considerations. Let us regard the double integral

$$\iint\limits_G f(x, y)\, dx\, dy$$

as the volume of a curvilinear cylinder T bounded below by a domain G, above by a surface $z = f(x, y)$ and on the sides by a cylindrical

surface passing through the boundary of the domain G (see Fig. 1.13). The solid T can be thought of as being composed of infinitely thin layers parallel to the y, z-plane. The volume of each layer is equal to the product

$$J(x)\,dx$$

where $J(x)$ is the area of the corresponding section of the solid T and dx is the width of the layer. Then the total volume of the solid T is equal to

$$\int_a^b J(x)\,dx \tag{1.33}$$

But the area $J(x)$ (as the area of a curvilinear trapezoid) is given by the integral

$$\int_{y_1(x)}^{y_2(x)} f(x, y)\,dy \tag{1.34}$$

where x is regarded as a fixed quantity and the quantities $y_1(x)$ and $y_2(x)$ are the coordinates of the end points of the line segment

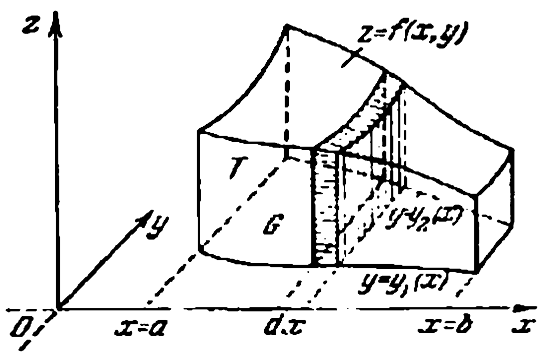


Fig. 1.13

which is the projection of the section on the x, y-plane. Combining (1.33) and (1.34) we see that the volume of the solid T can be expressed in the form

$$\int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy$$

Hence, we obtain the relation

$$\iint_G f(x, y)\,ds = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy \tag{1.35}$$

This formula tells us that when a double integral is thought of as a sum of the elements $f(x, y)\,dx\,dy$ we can first perform the summation within the layers parallel to one of the coordinate planes and then sum up the results corresponding to each layer. As an algebraic

analogue of relation (1.35), we can write the well known formula

$$\sum_{i,k} a_{ik} = \sum_i \left(\sum_k a_{ik}\right)$$

It is clear that if we took the sections of a curvilinear cylinder parallel to the x, z-plane instead of the y, z-plane this would result in the equality

$$\iint_G f(x, y)\, ds = \int_c^d dy \int_{x_1(y)}^{x_2(y)} f(x, y)\, dx$$

(see Fig. 1.14). Now let us pass from our heuristic considerations to strict arguments.

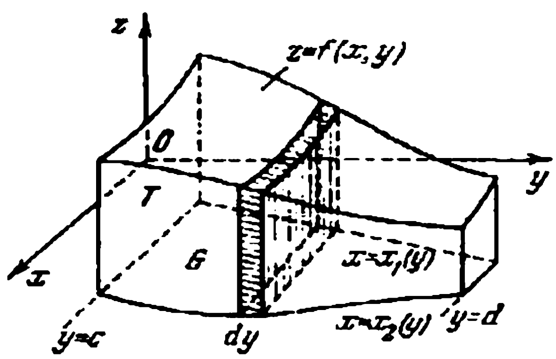


Fig. 1.14

2. The Case of a Rectangular Domain of Integration. We begin with a double integral taken over a rectangle with sides parallel to the coordinate axes.

Theorem 1.5. If $f(x, y)$ is a function defined in the rectangle

$$P = \{a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d\} \tag{1.36}$$

for which the double integral

$$\iint_P f(x, y)\, dx\, dy \tag{1.37}$$

exists and if the onefold (single) integral

$$J(x) = \int_c^d f(x, y)\, dy \tag{1.38}$$

exists for each fixed value of x in the interval $a \leqslant x \leqslant b$ then the ***iterated (repeated) integral***

$$\int_a^b dx \int_c^d f(x, y)\, dy = \int_a^b J(x)\, dx \tag{1.39}$$

also exists, and we have

$$\iint_P f(x, y)\, dx\, dy = \int_a^b dx \int_c^d f(x, y)\, dy \tag{1.40}$$

Proof. Divide the rectangle P into rectangular subdomains P_{ij} by breaking up its sides with the help of points $a = x_0 < x_1 < < x_2 < \ldots < x_k = b$ and, respectively, $c = y_0 < y_1 < y_2 < \ldots \ldots < y_l = d$. Thus, P_{ij} is the rectangle of the form $P_{ij} = \{x_{i-1} \leqslant x \leqslant x_i,\ y_{j-1} \leqslant y \leqslant y_j\}$ (see Fig. 1.15). Let m_{ij} be

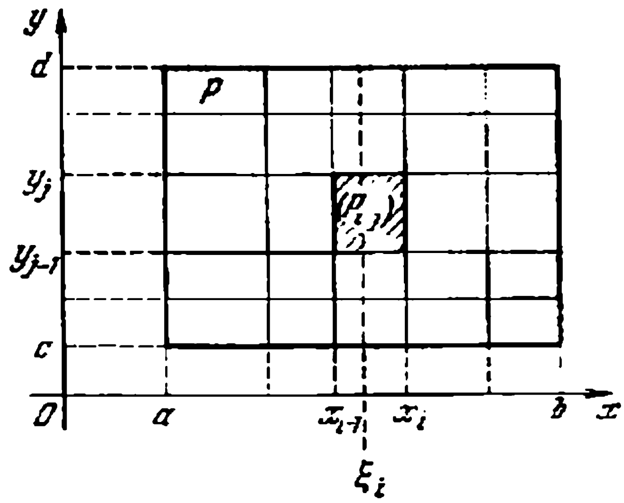


Fig. 1.15

the greatest lower bound and M_{ij} the least upper bound, on the rectangle P_{ij}, of the values of the function $f(x, y)$. Choose a point ξ_i in each subinterval $[x_{i-1}, x_i]$. Since $m_{ij} \leqslant f(\xi_i, y) \leqslant M_{ij}$ for $y_{j-1} \leqslant y \leqslant y_j$, we have

$$m_{ij}\Delta y_j \leqslant \int\limits_{y_{j-1}}^{y_j} f(\xi_i, y)\,dy \leqslant M_{ij}\Delta y_j \quad (\Delta y_j = y_j - y_{j-1}) \qquad (1.41)$$

and the integral in (1.41) exists because, according to the hypothesis, integral (1.38) taken over the whole interval $[c, d]$ exists for every x. Summing up inequalities (1.41) with respect to j from 1 to l we derive

$$\sum_{j=1}^{l} m_{ij}\Delta y_j \leqslant J(\xi_i) = \int\limits_{c}^{d} f(\xi_i, y)\,dy \leqslant \sum_{j=1}^{l} M_{ij}\Delta y_j, \quad i = 1, 2, \ldots, k$$

Multiplying each of the last inequalities by $\Delta x_i = x_i - x_{i-1}$ and summing them with respect to i from 1 to k we deduce

$$\sum_{i=1}^{k} \Delta x_i \sum_{j=1}^{l} m_{ij}\Delta y_j \leqslant \sum_{i=1}^{k} J(\xi_i)\,\Delta x_i \leqslant \sum_{i=1}^{k} \Delta x_i \sum_{j=1}^{l} M_{ij}\Delta y_j$$

The expression $\sum\limits_{i=1}^{k} J(\xi_i)\,\Delta x_i$ entering into this relation is an integral sum associated with the function $J(x)$ whereas $\sum\limits_{i=1}^{k} \Delta x_i \sum\limits_{j=1}^{l} m_{ij}\Delta y_j$ and $\sum\limits_{i=1}^{k} \Delta x_i \sum\limits_{j=1}^{l} M_{ij}\Delta y_j$ are the lower and

the upper Darboux sums corresponding to double integral (1.37). Consequently,

$$\omega \leqslant \sum_{i=1}^{k} J(\xi_i)\, \Delta x_i \leqslant \Omega$$

If we now make all Δx_i and Δy_j tend to zero then, since we have supposed that double integral (1.37) exists,* both the lower and the upper sums will tend to the double integral. Hence, the integral sums $\sum_{i=1}^{k} J(\xi_i)\, \Delta x_i$ tend to the same limit. Thus, we have

$$\iint_P f(x, y)\, dx\, dy = \int_a^b J(x)\, dx = \int_a^b dx \int_c^d f(x, y)\, dy$$

Interchanging the variables x and y (and supposing that the integral $J_1(y) = \int_a^b f(x, y)\, dx$ exists) we derive a similar relation of the form

$$\int_c^d dy \int_a^b f(x, y)\, dx = \iint_P f(x, y)\, dx\, dy$$

Finally, if both integrals $J(x) = \int_c^d f(x, y)\, dy$ and $J_1(y) = \int_a^b f(x, y)\, dx$ exist, together with integral (1.37), we obtain

$$\iint_P f(x, y)\, dx\, dy = \int_a^b dx \int_c^d f(x, y)\, dy = \int_c^d dy \int_a^b f(x, y)\, dx$$

3. The Case of a Curvilinear Domain. We now pass to the question on reducing a double integral to an iterated one for the case of a curvilinear domain. Let a domain G be bounded by two continuous curves $y = y_1(x)$ and $y = y_2(x)$ (where $y_2(x) \geqslant y_1(x)$) and by vertical line segments $x = a$ and $x = b$ (Fig. 1.16). Then the follow-

---

* By the hypothesis, double integral (1.37) exists and therefore, for any way of partitioning the rectangle P into subdomains such that their maximal diameter tends to zero, the upper and the lower Darboux sums tend to the common limit, i.e. to the corresponding double integral. This enables us to realize the partition in any appropriate manner, and we have chosen the one performed by means of vertical and horizontal straight lines.

ing theorem takes place:

Theorem 1.6. *If the double integral*

$$\iint\limits_G f(x, y)\,dx\,dy$$

exists for a function $f(x, y)$ defined in the domain G, and the integral

$$J(x) = \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\,dy$$

exists for each fixed value of x from the interval $[a, b]$ the iterated integral

$$\int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\,dy$$

also exists and we have the equality

$$\iint\limits_G f(x, y)\,dx\,dy = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\,dy \tag{1.42}$$

Proof. Put $c = \min y_1(x)$, $d = \max y_2(x)$ and embed the domain G in the rectangle P determined by the inequalities $a \leqslant x \leqslant b$,

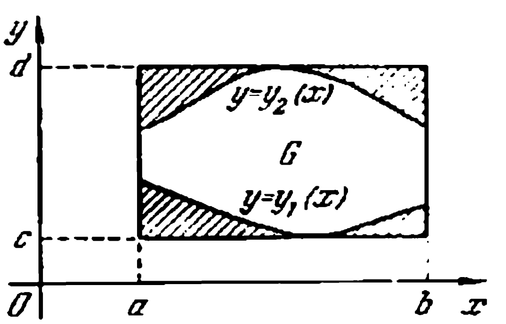


Fig. 1.16

$c \leqslant y \leqslant d$ (see Fig. 1.16). Consider the auxiliary function $f^*(x, y)$ defined on the rectangle by the relations

$$f^*(x, y) = \begin{cases} f(x, y) & \text{in } G \\ 0 & \text{in } P - G \end{cases}$$

where $P - G$ is the difference between the sets P and G, i.e. the collection of all points of P not belonging to G.

The function $f^*(x, y)$ satisfies the conditions of the foregoing theorem. In fact, since it coincides with $f(x, y)$ on the domain G it is integrable in G and it is identically equal to zero in $P - G$ and thus is integrable there too. Consequently, by the property of

additivity (see § 5, Sec. 4, property 3), the function is integrable over the entire rectangle P. Furthermore, we have

$$\iint_{G} f^*(x, y)\,dx\,dy = \iint_{G} f(x, y)\,dx\,dy$$

and

$$\iint_{P-G} f^*(x, y)\,dx\,dy = 0$$

whence

$$\iint_{P} f^*(x, y)\,dx\,dy = \iint_{G} f(x, y)\,dx\,dy \tag{1.43}$$

Besides, for each value of x lying between a and b, the integral

$$\int_c^d f^*(x, y)\,dy = \int_c^{y_1(x)} f^*(x, y)\,dy + \int_{y_1(x)}^{y_2(x)} f^*(x, y)\,dy +$$
$$+ \int_{y_2(x)}^{d} f^*(x, y)\,dy \tag{1.44}$$

is sure to exist because each of the three integrals entering into the right-hand side exists. Actually, the line segments connecting, respectively, the points (x, c), $(x, y_1(x))$ and $(x, y_2(x))$, (x, d) in the x, y-plane lie outside the domain G and $f^*(x, y)$ equals zero on them, and the integral $\int\limits_{y_1(x)}^{y_2(x)} f^*(x, y)\,dy$ coincides with the integral

$$\int_{y_1(x)}^{y_2(x)} f(x, y)\,dy$$

which exists by the hypothesis. The first and the third integrals entering into the right-hand side of (1.44) being equal to zero, we finally obtain

$$\int_c^d f^*(x, y)\,dy = \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy \tag{1.45}$$

We see that the function $f^*(x, y)$ defined in the rectangle P satisfies the conditions of Theorem 1.5 and, consequently, the double integral of it over P can be reduced to the iterated integral:

$$\iint_{P} f^*(x, y)\,dx\,dy = \int_a^b dx \int_c^d f^*(x, y)\,dy$$

From the last relation and equalities (1.43) and (1.45) we deduce

$$\iint_G f(x, y)\,dx\,dy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy$$

which is what we set out to prove.

In Theorem 1.6 we have considered a domain G such that every vertical straight line $x = \text{const}$ cuts its boundary at no more than two points $(x, y_1(x))$ and $(x, y_2(x))$ and supposed that the integral

$$J(x) = \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy \quad (a \leqslant x \leqslant b)$$

exists. If we suppose that every straight line $y = \text{const}$ has at most two common points $(x_1(y), y)$ and $(x_2(y), y)$ with the boundary of a domain G (see Fig. 1.17) and require that the integral $\int_{x_1(y)}^{x_2(y)} f(x, y)\,dx$ should exist for each fixed y we can prove the existence of the iterated integral

$$\int_c^d dy \int_{x_1(y)}^{x_2(y)} f(x, y)\,dx$$

and its coincidence with the double integral $\iint_G f(x, y)\,dx\,dy$.

As was seen at the beginning of § 5, the geometric meaning of the formulas reducing a double integral to an iterated one is

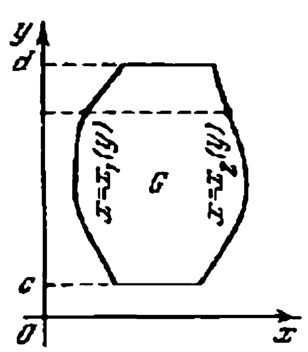


Fig. 1.17

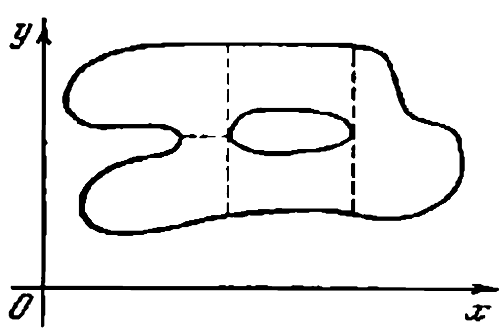


Fig. 1.18

that the volume of a solid is equal to the integral of the area of its cross section (which is a function of the variable determining the position of the cutting plane).

Note 1. If the domain G is such that there are straight lines (vertical or horizontal), passing through interior points of the domain, which have more than two common points with its boundary, then, to represent the double integral taken over the domain in the form of

an iterated integral, one should divide the domain G into parts satisfying the conditions of Theorem 1.6 and separately reduce each of the corresponding double integrals to an iterated one (see Fig. 1.18).

For example, let the domain of integration G be the unit circle $x^2+y^2\leqslant 1$ from which the ellipse $x^2+2y^2\leqslant 1$ is cut out

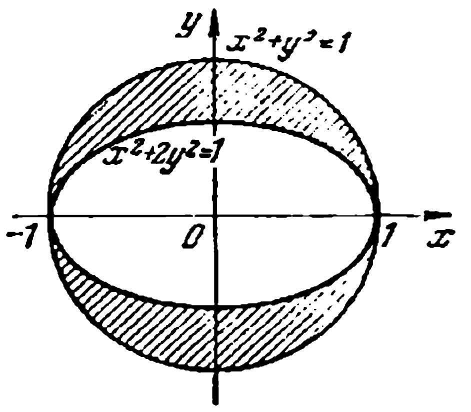


Fig. 1.19

(Fig. 1.19). Then the double integral over G can be, for instance represented as

$$\iint\limits_G f(x, y)\,dx\,dy=\int\limits_{-1}^{1} dx \int\limits_{\sqrt{\frac{1-x^2}{2}}}^{\sqrt{1-x^2}} f(x, y)\,dy+$$

$$+\int\limits_{-1}^{1} dx \int\limits_{-\sqrt{1-x^2}}^{-\sqrt{\frac{1-x^2}{2}}} f(x, y)\,dy$$

that is in the form of a sum of two iterated integrals.

Note 2. If a double integral can be reduced both to an iterated integral of the form $\int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\,dy$ and to an iterated integral of the form $\int\limits_c^d dy \int\limits_{x_1(y)}^{x_2(y)} f(x, y)\,dx$ then, when computing the double integral, we can use any of these representations. But it may well happen that one of them is more convenient than the other, and therefore, in concrete problems, an appropriate choice of the order of integration (i.e. the order in which integrations with respect to x and y are performed) may be of essential significance.

Exercise. Write the double integral

$$\iint\limits_G f(x, y)\,dx\,dy$$

where G is the domain bounded by the curves $y = \sqrt{2ax - x^2}$ and $y = \sqrt{2ax}$ and by the straight line $x = 2a$ (Fig. 1.20) in the form

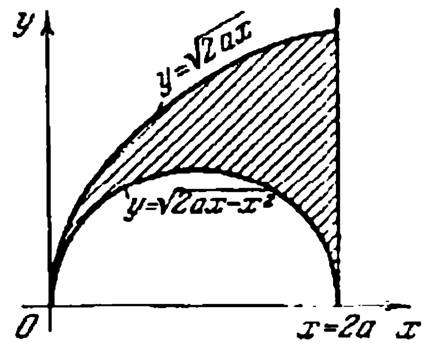


Fig. 1.20

of an iterated integral (for both possible orders of integration).

Answer. (1) $\int\limits_0^{2a} dx \int\limits_{\sqrt{2ax-x^2}}^{\sqrt{2ax}} f(x, y)\, dy$

(2) $$\int\limits_a^{2a} dy \int\limits_{\frac{y^2}{2a}}^{2a} f(x, y)\, dx + \int\limits_0^{a} dy \int\limits_{\frac{y^2}{2a}}^{a-\sqrt{a^2-y^2}} f(x, y)\, dx + \int\limits_0^{a} dy \int\limits_{a+\sqrt{a^2-y^2}}^{2a} f(x, y)\, dx$$

In the second case we have to break the integral into three summands whereas the first case involves only one term.

§ 6. CHANGE OF VARIABLES IN DOUBLE INTEGRAL

We often apply a change of variables when we integrate a function of one independent variable, and this method is also very important for evaluating double integrals. Before studying the problem of changing variables in a double integral we shall discuss some questions related to mappings of domains.

1. Mapping of Plane Figures. Consider two planes with respective Cartesian coordinates x, y and ξ, η in them. Suppose that in the x, y-plane we have a bounded closed domain G with boundary L and in the ξ, η-plane a bounded closed domain Γ* (see Fig. 1.21a and b). Let

$$x = x(\xi, \eta), \quad y = y(\xi, \eta) \tag{1.46}$$

be two functions defined in the domain Γ. Suppose that when the point (ξ, η) runs over the domain Γ the corresponding point (x, y)

* As before, we suppose that the domains G and Γ are squarable

runs over the domain G. Thus, functions (1.46) define a **mapping of the domain Γ onto the domain G.**

Let the mapping satisfy the following conditions:

(1) The mapping is one-to-one, which means that, to distinct points of the domain Γ, there correspond distinct points of the domain G. In other words, the solutions

$$\xi = \xi(x, y), \quad \eta = \eta(x, y) \tag{1.47}$$

of equations (1.46) (obtained by resolving the equations in ξ and η) are uniquely defined throughout the whole domain G.

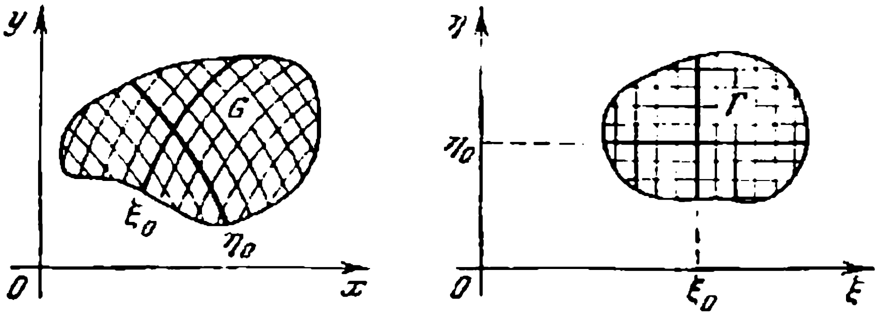


Fig. 1.21 (a) (b)

(2) Functions (1.46) and (1.47) are continuous and possess continuous partial derivatives of the first order.

(3) The *functional determinant* (*Jacobian*)

$$\frac{D(x, y)}{D(\xi, \eta)} = \begin{vmatrix} \dfrac{\partial x}{\partial \xi} & \dfrac{\partial x}{\partial \eta} \\ \dfrac{\partial y}{\partial \xi} & \dfrac{\partial y}{\partial \eta} \end{vmatrix} \tag{1.48}$$

is different from zero everywhere in the domain Γ, and, consequently, since the derivatives entering into the Jacobian are supposed to be continuous, it retains its sign in Γ.

The Jacobian $\dfrac{D(\xi, \eta)}{D(x, y)}$ of inverse mapping (1.47) is connected with Jacobian (1.48) by the relation

$$\frac{D(x, y)}{D(\xi, \eta)} \cdot \frac{D(\xi, \eta)}{D(x, y)} = 1$$

which is directly implied by the definition of the product of determinants and the rules for differentiating a composite function. Therefore the Jacobian $\dfrac{D(\xi, \eta)}{D(x, y)}$ does not vanish in the domain G.

If we are given a smooth or piecewise smooth curve

$$\xi = \xi(t), \quad \eta = \eta(t), \quad \alpha \leqslant t \leqslant \beta$$

in the domain Γ, mapping (1.46) transforms it into the curve

$$x = x(\xi(t), \eta(t)) = x(t), \quad y = y(\xi(t), \eta(t)) = y(t)$$

which is again smooth or piecewise smooth. Indeed, the derivatives $\frac{d\xi}{dt}$ and $\frac{d\eta}{dt}$ existing and being continuous, the derivatives

$$\frac{dx}{dt}=\frac{\partial x}{\partial \xi}\frac{d\xi}{dt}+\frac{\partial x}{\partial \eta}\frac{d\eta}{dt} \text{ and } \frac{dy}{dt}=\frac{\partial y}{\partial \xi}\frac{d\xi}{dt}+\frac{\partial y}{\partial \eta}\frac{d\eta}{dt}$$

also exist and are continuous. Furthermore, if at least one of the derivatives $\frac{d\xi}{dt}$ and $\frac{d\eta}{dt}$ is different from zero the derivatives $\frac{dx}{dt}$ and $\frac{dy}{dt}$ do not vanish simultaneously since $\frac{D(x, y)}{D(\xi, \eta)} \neq 0$.

We can also assert that the boundary Λ of the domain Γ is mapped on the boundary L of the domain G. This follows from the theorem on implicit functions (e.g. see [8], Chapter 15, § 2). Indeed, if, to a point (x_0, y_0) belonging to L, an interior point (ξ_0, η_0) of the domain Γ corresponded, the relations

$$x = x(\xi, \eta), \quad y = y(\xi, \eta)$$

would define the quantities ξ and η as functions of x and y in a neighbourhood of the point (x_0, y_0). But every neighbourhood of a boundary point contains points not belonging to G, and hence the point (ξ_0, η_0), an interior point of Γ, would possess a neighbourhood lying in Γ and not mapped into G which contradicts the hypothesis.

2. Curvilinear Coordinates. Consider a straight line $\xi = \xi_0$ in the domain Γ (Fig. 1.21). It is mapped on a smooth curve lying in the domain G and determined by the parametric equations

$$x = x(\xi_0, \eta), \quad y = y(\xi_0, \eta) \tag{1.49}$$

(where η is the parameter). Similarly, to each straight line $\eta = \eta_0$ in Γ there corresponds a curve in the domain G with parametric equations

$$x = x(\xi, \eta_0), \quad y = y(\xi, \eta_0) \tag{1.50}$$

where ξ is the parameter. Curves (1.49) and (1.50), lying in the domain G, into which mapping (1.46) transforms straight lines parallel to the coordinate axes ξ, η and belonging to Γ, are referred to as the **coordinate curves** (ξ-**curves** and η-**curves**) in the domain G.

The mapping

$$x = x(\xi, \eta), \quad y = y(\xi, \eta)$$

being one-to-one, it follows that there is a single curve of form (1.49) passing through each point (x, y) of the domain G which corresponds to a given constant value of ξ and a single curve of form (1.50) corresponding to a constant value of η. Consequently, the quantities ξ and η can be regarded as the coordinates (different, of course, from the Cartesian ones) of points belonging to the domain G. The coordinate lines (1.49) and (1.50) corresponding to these

coordinates being curvilinear (but in the general case, not straight lines as in the case of Cartesian coordinates), the quantities ξ and η are called the **curvilinear coordinates** of the points of the domain G.

Thus, from the geometric point of view, the variables ξ and η are interpreted in a twofold sense: on the one hand, they are the Cartesian coordinates of the points of the domain Γ and, on the other hand, they are the curvilinear coordinates of the points belonging to the domain G. Accordingly, every relation of the form $\Phi(\xi, \eta) = 0$ can be regarded as an equation (in Cartesian coordinates) determining a curve λ lying in the domain Γ and also as an equation (in curvilinear coordinates) of a curve l, the image of the curve λ under mapping (1.46), placed in G.

3. Polar Coordinates. The connection between polar coordinates r, φ and Cartesian coordinates x, y is given by the relations

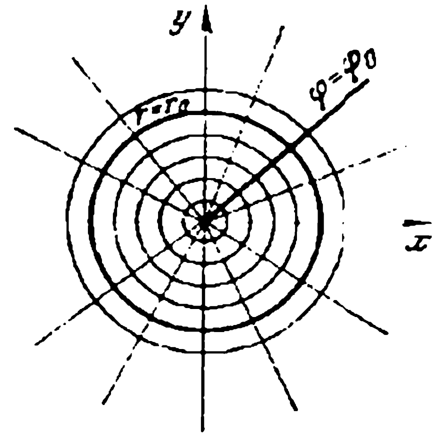
$$x = r\cos\varphi, \quad y = r\sin\varphi \ (r \geqslant 0,\ 0 \leqslant \varphi < 2\pi) \tag{1.51}$$

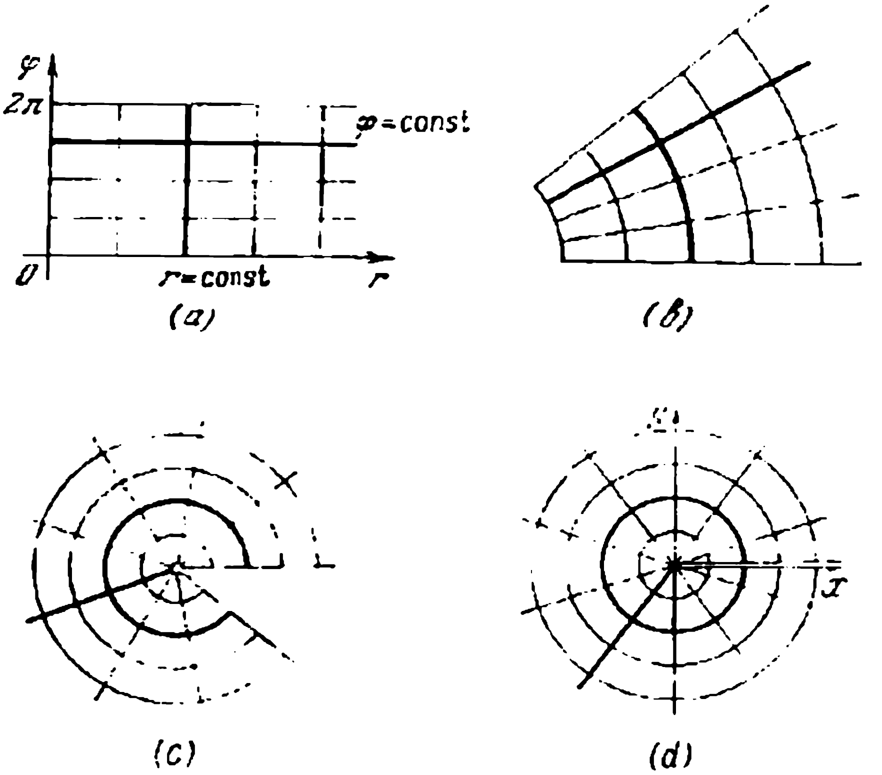
if the pole coincides with the origin.

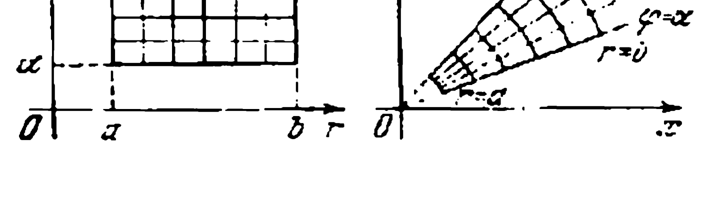
The coordinate curves of the polar system shown in Fig. 1.22 are the concentric circles with centre at the origin ($r = \text{const}$) and the rays starting from the centre ($\varphi = \text{const}$). Mapping (1.51) transforms the half-strip $r \geqslant 0$, $0 \leqslant \varphi < 2\pi$ onto the entire x, y-plane. The mapping is one-to-one everywhere except the point $x = 0$, $y = 0$ which is the image of the half-segment $r = 0$, $0 \leqslant \varphi < 2\pi$ of the r, φ-plane. If we delete the point $x = 0$, $y = 0$ we can consider the inverse of mapping (1.51) which transforms the punctured x, y-plane into the half-strip $r > 0$, $0 \leqslant \varphi < 2\pi$. The inverse mapping is continuous everywhere except the positive half of x-axis because, although the value $\varphi = 0$ corresponds to all the points of the semiaxis, the variable φ tends to 2π but not to zero as the point $M(x, y)$ approaches the semiaxis from below. Thus, formulas (1.51) define the mapping of the half-strip $r \geqslant 0$, $0 \leqslant \varphi < 2\pi$ onto the x, y-plane which is the one-to-one and continuous, together with its inverse mapping, everywhere except the points at which $r = 0$ or $\varphi = 0$.

We can visualize the transformation from the half-strip on the r, φ-plane to the x, y-plane as "spreading a fan". Imagine that we are watching a film showing the process of spreading out the half-strip $0 \leqslant r < \infty$, $0 \leqslant \varphi < 2\pi$ as a fan which, when opened, covers the x, y-plane. Fig. 1.23*a* represents the first still of the film, Fig. 1.23*b* the second, Fig. 1.23*c* shows one of the final stills and Fig. 1.23*d* the last one.

Consider an example. Let a rectangular domain in the r, φ-plane be given by the relations $0 < a \leqslant r \leqslant b$, $0 \leqslant \alpha \leqslant \varphi \leqslant \beta < 2\pi$. The "fan spreading" procedure transforms it into an annular sector in the x, y-plane (Fig. 1.24).


y
x
r=r0
φ=φ0


φ
2π
φ=const
0
r=const
r
(a)
(b)
(c)
(d)
x


φ
β
α
0
a
b
r
y
φ=β
φ=α
r=b
r=a
0
x

Let us find the Jacobian corresponding to the transformation from Cartesian coordinates to polar ones, i.e. the Jacobian of functions (1.51). We obtain

$$\frac{D(x, y)}{D(\xi, \eta)}=\begin{vmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{vmatrix}=r$$

The Jacobian is different from zero everywhere except the point $x_1 = 0$, $y = 0$.

4. Statement of the Problem of Changing Variables in the Double Integral. We now formulate the general problem of changing variables in the double integral. Let G be a closed domain bounded by a piecewise smooth curve L and $f(x, y)$ a bounded function, defined in G, which is continuous or has discontinuities forming a set of area zero. Further, let the functions

$$x = x(\xi, \eta), \quad y = y(\xi, \eta)$$

determine a mapping of a domain Γ on the domain G and let the mapping satisfy the conditions (1)-(3) enumerated in Sec. 1. Our aim is to represent the integral

$$\iint\limits_G f(x, y)\,dx\,dy$$

taken over the domain G as an integral over the domain Γ by transforming its element of integration to the new variables ξ and η.

5. Computing Area in Curvilinear Coordinates. In deriving the formula for changing variables in a double integral the chief step is to express the area of a domain in curvilinear coordinates. The following theorem solves the problem:

Theorem 1.7. Let $x = x(\xi, \eta)$, $y = y(\xi, \eta)$ be a continuous and continuously differentiable one-to-one mapping of a domain Γ of the ξ, η-plane on a domain G in the x, y-plane. Suppose that the Jacobian $\frac{D(x, y)}{D(\xi, \eta)}$ is everywhere different from zero. Then

$$\text{area of } G=\iint\limits_G dx\,dy=\iint\limits_\Gamma\left|\frac{D(x, y)}{D(\xi, \eta)}\right|d\xi\,d\eta \tag{1.52}$$

We shall preface the proof of the theorem with an intuitive argument (to which, by the way, the reader may restrict himself).

Consider two pairs of coordinate curves lying infinitely close to each other in the domain G. Let the first pair correspond to the values

$$\xi_0 \quad\text{and}\quad \xi_0 + d\xi$$

of the coordinate ξ and the second pair to the values

$$\eta_0 \quad\text{and}\quad \eta_0 + d\eta$$

of the coordinate η. These coordinate curves cut out of the domain G an infinitesimal element of area $A_0A_1A_3A_2$ (see Fig. 1.25). This element can be apparently regarded as a parallelogram to within

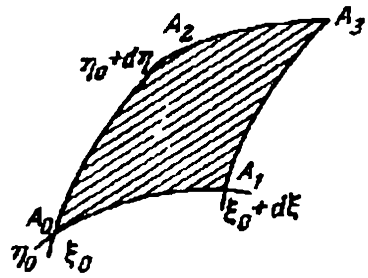


Fig. 1.25

infinitesimals of order higher than the first. We see that the sides of the parallelogram are the vectors

$$\overline{A_0A_1}=\left(\frac{\partial x}{\partial \xi}\,d\xi,\ \frac{\partial y}{\partial \xi}\,d\xi\right)$$

and

$$\overline{A_0A_2}=\left(\frac{\partial x}{\partial \eta}\,d\eta,\ \frac{\partial y}{\partial \eta}\,d\eta\right)$$

The area ds of the parallelogram $A_0A_1A_3A_2$ is equal to the modulus of the determinant whose elements are the projections of the vectors $\overline{A_0A_1}$ and $\overline{A_0A_2}$ on the coordinate axes, that is

$$ds=\text{modulus of}\begin{vmatrix}\frac{\partial x}{\partial \xi}\,d\xi & \frac{\partial y}{\partial \xi}\,d\xi\\ \frac{\partial x}{\partial \eta}\,d\eta & \frac{\partial y}{\partial \eta}\,d\eta\end{vmatrix}=\left|\frac{D(x,y)}{D(\xi,\eta)}\right|d\xi\,d\eta \tag{1.53}$$

Hence, the total area S of the domain G is obtained by summing up all the elements, i.e. is in fact representable in the form of a double integral taken over the range Γ of the variables ξ and η:

$$S=\iint_{\Gamma}\left|\frac{D(x,y)}{D(\xi,\eta)}\right|d\xi\,d\eta$$

We can now corroborate our intuitive argument with a proof. We shall leave out some details which we believe will not be difficult for the reader to understand. Besides, to simplify the consideration we shall suppose that the mapping is defined and satisfies the requirements of the theorem not only in the domain Γ but also in a wider domain in which Γ is strictly contained together with its boundary.

Proof of Theorem 1.7. We first take a simple but fundamental case when the domain in the ξ, η-plane is a rectangle Π with sides parallel to the coordinate axes, and the mapping of the rectangle

on the x, y-plane is a *linear* one expressed by the formulas

$$x = x_0 + a\xi + b\eta, \quad y = y_0 + a_1\xi + b_1\eta \tag{1.54}$$

where x_0, y_0, a, b, a_1 and b_1 are constants and $\begin{vmatrix} a & b \\ a_1 & b_1 \end{vmatrix} \neq 0$.

As is well known from analytic geometry, the image of such a rectangle Π under a mapping of this type is a parallelogram P whose area is connected with the area of the rectangle Π by the relation

$$\text{area of } P = \left(\text{modulus of } \begin{vmatrix} a & b \\ a_1 & b_1 \end{vmatrix}\right) \cdot (\text{area of } \Pi) \tag{1.55}$$

(prove it). It follows that every squarable figure Φ lying in the ξ, η-plane is mapped by linear mapping (1.54) onto a squarable figure F in the x, y-plane whose area is expressed as

$$\text{area of } F = \left(\text{modulus of } \begin{vmatrix} a & b \\ a_1 & b_1 \end{vmatrix}\right) \cdot (\text{area of } \Phi) \tag{1.56}$$

By the way, it is only relation (1.55) that we need for our further aims.

Now let us consider an arbitrary (possibly nonlinear) mapping $x = x(\xi, \eta)$, $y = y(\xi, \eta)$, of a domain Γ, satisfying the conditions

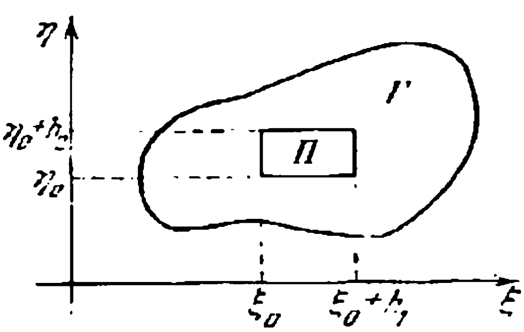


Fig. 1.26

of the theorem. Take a point (ξ_0, η_0) belonging to the domain Γ where the mapping is defined and consider the rectangle

$$\xi_0 \leqslant \xi \leqslant \xi_0 + h_1, \quad \eta_0 \leqslant \eta \leqslant \eta_0 + h_2$$

which we again denote by Π (Fig. 1.26).

Applying Lagrange's theorem on finite increments we rewrite the equations defining the mapping of the rectangle into the x, y-plane in the form

$$x = x_0 + \frac{\partial x}{\partial \xi}\,d\xi + \frac{\partial x}{\partial \eta}\,d\eta + \alpha_1, \quad y = y_0 + \frac{\partial y}{\partial \xi}\,d\xi + \frac{\partial y}{\partial \eta}\,d\eta + \alpha_2 \tag{1.57}$$

where $x_0 = x(\xi_0, \eta_0)$, $y_0 = y(\xi_0, \eta_0)$, the values of the derivatives are taken at the point (ξ_0, η_0) and

$$\alpha_1 = (x'_\xi(\xi^*, \eta^*) - x'_\xi(\xi_0, \eta_0))\,d\xi + (x'_\eta(\xi^*, \eta^*) - x'_\eta(\xi_0, \eta_0))\,d\eta$$
$$\alpha_2 = (y'_\xi(\xi^{**}, \eta^{**}) - y'_\xi(\xi_0, \eta_0))\,d\xi + (y'_\eta(\xi^{**}, \eta^{**}) - y'_\eta(\xi_0, \eta_0))\,d\eta$$

(Here we have $\xi_0 \leqslant \xi^* \leqslant \xi$; $\xi_0 \leqslant \xi^{**} \leqslant \xi$; $\eta_0 \leqslant \eta^* \leqslant \eta$; $\eta_0 \leqslant \eta^{**} \leqslant \eta$.)

By the hypothesis, the first derivatives of x and y with respect to ξ and η are continuous and thus they are uniformly continuous in the bounded closed domain Γ. Consequently, for every $\varepsilon > 0$ there exists a sufficiently small $h > 0$ such that if $h_1 + h_2 < h$ the inequalities

$$|x'_\xi(\xi, \eta) - x'_\xi(\xi_0, \eta_0)| < \varepsilon, \quad |x'_\eta(\xi, \eta) - x'_\eta(\xi_0, \eta_0)| < \varepsilon$$

hold for all the points (ξ, η) belonging to the rectangle Π, and similar inequalities are true for y'_ξ and y'_η, ε being independent of the particular choice of the point (ξ_0, η_0). These estimates show that

$$|\alpha_1| < \varepsilon h, \quad |\alpha_2| < \varepsilon h \tag{1.58}$$

Now compare nonlinear mapping (1.57) with the linear mapping

$$\hat{x} = x_0 + \frac{\partial x}{\partial \xi} d\xi + \frac{\partial x}{\partial \eta} d\eta, \quad \hat{y} = y_0 + \frac{\partial y}{\partial \xi} d\xi + \frac{\partial y}{\partial \eta} d\eta \tag{1.59}$$

which is obtained by dropping α_1 and α_2 in formulas (1.57). As we already know, a linear transformation of this kind maps the rectangle Π on a parallelogram which we again denote by P and, according to (1.55), we have

$$\text{area of } P = \left(\text{modulus of } \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{vmatrix}\right) \cdot (\text{area of } \Pi) \tag{1.60}$$

Nonlinear mapping (1.57) transforms Π into a curvilinear figure $\mathcal{P}$. Let us investigate what is the difference between its area and the area of the parallelogram P.

By virtue of (1.58), for each point $(\xi, \eta) \in \Pi$ we have

$$|x - \hat{x}| = |\alpha_1| < \varepsilon h, \quad |y - \hat{y}| = |\alpha_2| < \varepsilon h$$

i.e.

$$\sqrt{(x - \hat{x})^2 + (y - \hat{y})^2} < \sqrt{2}\, \varepsilon h$$

In other words, the distances between the images of the point $(\xi, \eta) \in \Pi$ under linear mapping (1.59) and nonlinear mapping (1.57) are less than $\sqrt{2}\,\varepsilon h$. Therefore, if we embed the boundary of the parallelogram P in a strip of width $\sqrt{2}\,\varepsilon h$ the boundary of the curvilinear figure $\mathcal{P}$ will be strictly contained in the strip (Fig. 1.27). It is clear that the difference between the area of $\mathcal{P}$ and the area of P does not exceed the area of the strip. Performing simple computations we find that the area of the strip is not greater than its width multiplied by the perimeter of the parallelogram P. The perimeter can be easily estimated. Let a positive number M be so chosen that each of the derivatives $\frac{\partial x}{\partial \xi}$, $\frac{\partial x}{\partial \eta}$, $\frac{\partial y}{\partial \xi}$ and $\frac{\partial y}{\partial \eta}$ does not

exceed M in its absolute value throughout the domain Γ (the derivatives are continuous and therefore bounded in the bounded closed domain Γ). Then (1.59) immediately implies that the sides of the parallelogram P cannot exceed Mh. Thus, the perimeter of P is

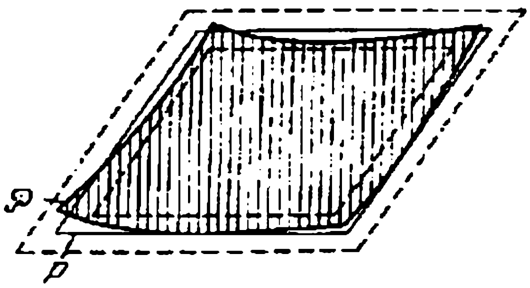


Fig. 1.27

not greater than $4Mh$, and the area of the strip in which the boundary of P has been embedded does not exceed $4\sqrt{2}\,\varepsilon Mh^2$, i.e. is not greater than

$$\sqrt{2}\,\varepsilon M\cdot(\text{area of }\Pi)$$

Consequently,

$$\text{area of }\mathscr{P}=\text{area of }P+\gamma$$

or, by virtue of (1.60), we have

$$\text{area of }\mathscr{P}=\left(\text{modulus of }\begin{vmatrix}\frac{\partial x}{\partial\xi} & \frac{\partial x}{\partial\eta}\\ \frac{\partial y}{\partial\xi} & \frac{\partial y}{\partial\eta}\end{vmatrix}\right)\cdot(\text{area of }\Pi)+\gamma \qquad (1.61)$$

where

$$|\gamma|<\sqrt{2}\,\varepsilon M\cdot(\text{area of }\Pi) \qquad (1.62)$$

Let now Φ be a polygonal figure lying within Γ and composed of rectangles with sides parallel to the coordinate axes, and $\mathscr{F}$ be its image under the mapping $x = x(\xi, \eta)$, $y = y(\xi, \eta)$. Break up Φ into rectangles Π_i so that the half-perimeter of each of them should be less than h. The union of the images $\mathscr{P}_i$ of the rectangles is the figure $\mathscr{F}$, and the area of each $\mathscr{P}_i$ can be represented in the form

$$\text{area of }\mathscr{P}_i=\left|\frac{D(x, y)}{D(\xi, \eta)}\right|_{\substack{\xi=\xi_i\\ \eta=\eta_i}}\cdot(\text{area of }\Pi_i)+\gamma_i \qquad (1.63)$$

where (ξ_i, η_i) is a point belonging to the rectangle Π_i and

$$|\gamma_i|<\sqrt{2}\,\varepsilon M\cdot(\text{area of }\Pi_i)$$

Sum up equalities (1.63) over all the rectangles Π_i. This yields

$$\sum_{i=1}^{n} (\text{area of } \mathcal{F}_i) = \text{area of } \mathcal{F} =$$

$$= \sum_{i=1}^{n} \left| \frac{D(x, y)}{D(\xi, \eta)} \right|_{\substack{\xi=\xi_i \\ \eta=\eta_i}} \cdot (\text{area of } \Pi_i) + \sum_{i=1}^{n} \gamma_i \qquad (1.64)$$

The first summand on the right-hand side of equality (1.64) is obviously an integral sum associated with the integral

$$\iint_{\Phi} \left| \frac{D(x, y)}{D(\xi, \eta)} \right| d\xi\, d\eta \qquad (1.65)$$

and the second one does not exceed the expression

$$\sqrt{2} M\varepsilon \sum_{i=1}^{n} (\text{area of } \Pi_i) = \sqrt{2} M\varepsilon \cdot (\text{area of } \Phi)$$

where ε can be made arbitrarily small (because the diameter of the partition of the figure Φ can be chosen as small as desired). The integrand being continuous, integral (1.65) is sure to exist. Consequently, we can pass to the limit in relation (1.64) as the partition of the figure Φ is infinitely refined. We thus obtain

$$\text{area of } \mathcal{F} = \iint_{\Phi} \left| \frac{D(x, y)}{D(\xi, \eta)} \right| d\xi\, d\eta$$

To complete the proof we should pass from the polygonal figure Φ embedded in the domain Γ to the domain Γ itself. This can be easily performed. The domain Γ being squarable, we can find two figures Φ_1 and Φ_2 composed of rectangles,* the first of which is embedded in Γ whereas the second envelops Γ, such that the difference between their areas is less than a given positive number δ. The mapping $x = x(\xi, \eta)$, $y = y(\xi, \eta)$ transforms them into two squarable figures $\mathcal{F}_1$ and $\mathcal{F}_2$, the former being embedded in G and the latter enveloping G. We can easily show that

$$|\text{area of } \mathcal{F}_1 - \text{area of } \mathcal{F}_2| < (2M^2 + \sqrt{2} M\varepsilon)\delta$$

(prove it by applying relation (1.64) and the inequality $\max \left| \frac{D(x, y)}{D(\xi, \eta)} \right| < 2M^2$). But then

$$|\text{area of } G - \text{area of } \mathcal{F}_1| < (2M^2 + \sqrt{2} M\varepsilon)\delta \qquad (1.66)$$

---

* The enveloping figure Φ_2 must lie in the domain in which Γ is strictly contained and in which, by the hypothesis, the mapping is defined and satisfies the conditions of the theorem.

The figure $\mathcal{F}_1$ is the image of the polygonal figure Φ_1 and therefore, by what has been proved, we have

$$\text{area of } \mathcal{F}_1 = \iint\limits_{\Phi_1} \left| \frac{D(x, y)}{D(\xi, \eta)} \right| d\xi\, d\eta \tag{1.67}$$

Furthermore, by the mean value theorem, we can write

$$\left| \iint\limits_{\Gamma} \left| \frac{D(x, y)}{D(\xi, \eta)} \right| d\xi\, d\eta - \iint\limits_{\Phi_1} \left| \frac{D(x, y)}{D(\xi, \eta)} \right| d\xi\, d\eta \right| =$$
$$= \iint\limits_{\Gamma - \Phi_1} \left| \frac{D(x, y)}{D(\xi, \eta)} \right| d\xi\, d\eta < 2M^2\delta \tag{1.68}$$

From (1.66) and (1.68), taking into account (1.67), we deduce

$$\left| \text{area of } G - \iint\limits_{\Gamma} \left| \frac{D(x, y)}{D(\xi, \eta)} \right| d\xi\, d\eta \right| < (4M^2 + \sqrt{2}\, M\varepsilon)\,\delta$$

The number δ being arbitrarily small here, the proof of the theorem thus follows.

Note 1. The fundamental idea upon which both the above proof and the foregoing intuitive argument are based lies in the fact that a nonlinear mapping $x = x(\xi, \eta)$, $y = y(\xi, \eta)$, when considered in the small, can be approximated with a linear one, and the smaller the domain, the greater the accuracy. By the way, properly speaking, the substitution of a linear relation for a nonlinear one, considered in the small, is, in general, the basic idea of mathematical analysis.

Example. Consider again polar coordinates. The curves $r = r_0$, $r = r_0 + dr$, $\varphi = \varphi_0$ and $\varphi = \varphi_0 + d\varphi$ cut out of the x, y-plane an infinitesimal rectangle with sides dr and $r_0\, d\varphi$ (see Fig. 1.28). Therefore the element of area in polar coordinates is equal to $r_0\, d\varphi\, dr$. (The same result is, of course, implied by general formula (1.52) since $\frac{D(x, y)}{D(r, \varphi)} = r$.) Consequently, the area in polar coordinates is expressed by the formula

$$S = \iint\limits_{\Gamma} r\, dr\, d\varphi \tag{1.69}$$

where Γ is the range of the variables r and φ. In particular, if the domain G is bounded by two rays $\varphi = \varphi_1$ and $\varphi = \varphi_2$ and by a curve $r = r(\varphi)$, i.e. has the form shown in Fig. 1.29 (represent the corresponding domain Γ in the r, φ-plane), then, reducing double integral (1.69) to an iterated integral, we obtain

$$S = \int\limits_{\varphi_1}^{\varphi_2} d\varphi \int\limits_{0}^{r(\varphi)} r\, dr$$

Performing integration with respect to r we find:

$$S=\frac{1}{2}\int\limits_{\varphi_1}^{\varphi_2} r^2(\varphi)\,d\varphi$$

This is the well known formula for area in polar coordinates (e.g. see [8], Chapter 11, § 2).

Note 2. Formula (1.53) indicates the geometric meaning of the absolute value of the Jacobian $\frac{D(y, x)}{D(\xi, \eta)}$. Actually, denote the

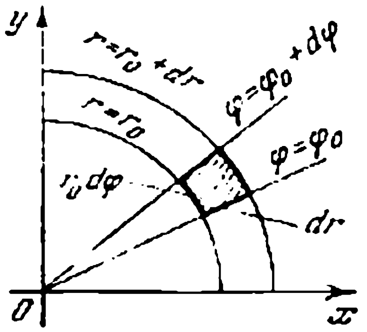


Fig. 1.28

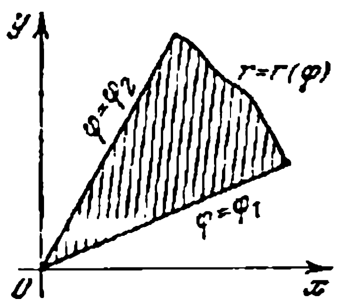


Fig. 1.29

Jacobian, for brevity, as $J(\xi, \eta)$ and consider the mapping of the domain Γ on the domain G determined by the formulas

$$x = x(\xi, \eta), \quad y = y(\xi, \eta)$$

The mapping transforms the infinitesimal rectangle belonging to Γ (see Fig. 1.30), which is bounded by the straight lines

$$\xi = \xi_0, \quad \xi = \xi_0 + d\xi \quad \text{and} \quad \eta = \eta_0, \quad \eta = \eta_0 + d\eta$$

and has the area $d\xi\, d\eta$, into a parallelogram of area

$$|J(\xi, \eta)|\, d\xi\, d\eta$$

Thus, the quantity $|J(\xi, \eta)|$ is the *coefficient of area expansion* (*at the point* (ξ, η)) for the mapping of the domain Γ on the domain G.

Note 3. In Theorem 1.7 we have supposed that the mapping

$$x = x(\xi, \eta), \quad y = y(\xi, \eta)$$

of the domain Γ onto domain G is one-to-one. But expression (1.52) for area in curvilinear coordinates remains true even when the condition is violated at some separate points or on separate curves. As a typical example of this kind, let us take the mapping of the rectangle $0 \leqslant r \leqslant a$, $0 \leqslant \varphi \leqslant 2\pi$ on the circle $x^2 + y^2 \leqslant a^2$, determined by the formulas

$$x = r\cos\varphi, \quad y = r\sin\varphi \tag{1.70}$$

which corresponds to the introduction of polar coordinates. The mapping satisfies the conditions of Theorem 1.7 everywhere except

the points belonging to the line segment $y = 0$, $0 \leqslant x \leqslant a$. Consider the rectangle $\rho \leqslant r \leqslant a$, $0 \leqslant \varphi \leqslant 2\pi - \varepsilon$ in the r, φ-plane

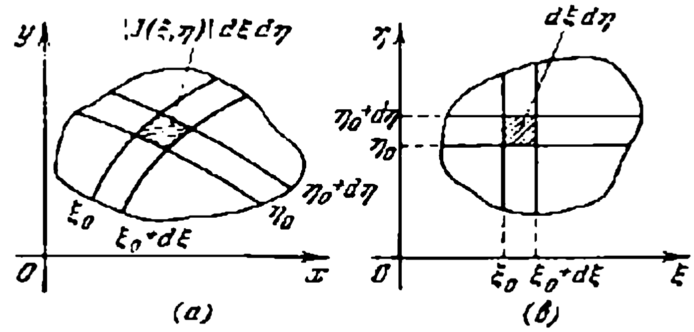


Fig. 1.30

(where $0 < \rho < a$ and $0 < \varepsilon < 2\pi$) and its image under mapping (1.70) in the x, y-plane (see Fig. 1.31). Formula (1.52) holds here because conditions (1)-(3) are fulfilled for these domains. Now,

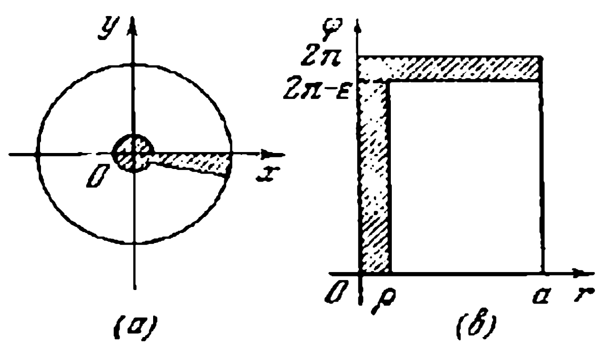


Fig. 1.31

passing to the limit, as $\rho \to 0$ and $\varepsilon \to 0$, we see that formula (1.52) remains true for the entire circle $r \leqslant a$.

Similar arguments are applicable in the general case of an arbitrary mapping which is one-to-one everywhere except separate points or curves.

6. Change of Variables in Double Integral. Expression (1.52) obtained for area in curvilinear coordinates enables us to derive the general formula for changing variables in a double integral. Consider the integral

$$\iint\limits_G f(x, y)\,dx\,dy \tag{1.71}$$

where the domain G is bounded by a piecewise smooth contour L and the function $f(x, y)$ is continuous throughout the domain (including its boundary) or is bounded and continuous in it everywhere possibly except a set of area zero.

Let functions $x = x(\xi, \eta)$ and $y = y(\xi, \eta)$ define a one-to-one correspondence between the points of the domain G and a domain Γ, and let the mapping satisfy all the requirements under which the validity of formula (1.52) expressing the area of the domain G in

curvilinear coordinates has been established. Divide the domain Γ into parts Γ_i by means of a system of piecewise smooth curves. The corresponding piecewise smooth curves in the x, y-plane (the images) break the domain G into parts G_i of area ΔS_i. Choose an arbitrary point (x_i, y_i) in each part G_i and form the integral sum

$$\sum_{i=1}^{n} f(x_i, y_i)\,\Delta S_i \tag{1.72}$$

associated with integral (1.71).

Applying formula (1.52) to each subdomain G_i we obtain

$$\Delta S_i = \iint\limits_{\Gamma_i} \left| \frac{D(x, y)}{D(\xi, \eta)} \right| d\xi\, d\eta$$

Denoting the Jacobian $\frac{D(x, y)}{D(\xi, \eta)}$ as $J(\xi, \eta)$ and taking advantage of the mean value theorem we can write

$$\Delta S_i = |J(\xi_i^*, \eta_i^*)|\, \Delta\sigma_i$$

where $\Delta\sigma_i$ is the area of the subdomain Γ_i. Now we substitute the above expression for the quantity ΔS_i in integral sum (1.72) and arrive at the sum

$$\sum_{i=1}^{n} f(x_i, y_i)\, |J(\xi_i^*, \eta_i^*)|\, \Delta\sigma_i$$

The point (ξ_i^*, η_i^*) appears when we apply the mean value theorem, and hence its position, within the subdomain Γ_i, is preassigned by the properties of the function and the subdomain, and we cannot take it at pleasure. But the point (x_i, y_i), unlike (ξ_i^*, η_i^*), is chosen in the corresponding subdomain G_i quite arbitrarily. Therefore we can put

$$x_i = x(\xi_i^*, \eta_i^*), \quad y_i = y(\xi_i^*, \eta_i^*)$$

i.e. take, as (x_i, y_i), the point of the subdomain G_i corresponding to the point (ξ_i^*, η_i^*) of the subdomain Γ_i. Then the above integral sum turns into the expression

$$\sum_{i=1}^{n} f(x(\xi_i^*, \eta_i^*), \quad y(\xi_i^*, \eta_i^*))\, |J(\xi_i^*, \eta_i^*)|\, \Delta\sigma_i$$

which is nothing but an integral sum for the integral

$$\iint\limits_{\Gamma} f(x(\xi, \eta), \quad y(\xi, \eta))\, |J(\xi, \eta)|\, d\xi\, d\eta \tag{1.73}$$

The integrand being continuous in the domain Γ or bounded and continuous everywhere in Γ except the points belonging to a set

of area zero, the integral is sure to exist. If we now make the maximal diameter of the partition of the domain Γ into the parts Γ_i tend to zero, the diameters of the subdomains G_i will also tend to zero. In this limiting process, the integral sum under consideration must tend both to double integral (1.71) and to integral (1.73). Hence, the integrals are equal:

$$\iint_G f(x, y)\,dx\,dy = \iint_\Gamma f(x(\xi, \eta), \quad y(\xi, \eta))\,|J(\xi, \eta)|\,d\xi\,d\eta \qquad (1.74)$$

It is this formula that describes the general case of changing variables in the double integral.

Thus, if G is a bounded closed domain with a piecewise smooth boundary and a function $f(x, y)$ defined in the domain is continuous throughout the domain or bounded and continuous in it everywhere except a set of area zero, and if the formulas

$$x = x(\xi, \eta), \qquad y = y(\xi, \eta)$$

determine a correspondence between the points of the domain G and a domain Γ lying in the ξ, η-plane which satisfies conditions (1)-(3) of Sec. 1, we have formula (1.74) for changing variables in the double integral.

Relation (1.74) also remains true when the condition that the mapping of the domain G on the domain Γ is one-to-one, continuous and continuously differentiable is violated at separate points or on a finite number of curves with zero area.

As in the case of an ordinary definite integral of a function of one argument, the method of changing variables is one of the most powerful tools for reducing a double integral to a form appropriate for computing it. But it should be noted that in the case of two independent variables there appears a new feature. The matter is that a change of variable is introduced in a definite integral in order to simplify the element of integration whereas when changing variables in a double integral we try to simplify not only the integrand but also the shape of the domain of integration. What has been said is so important that it is sometimes advisable to complicate the integrand when this yields a simplification of the domain of integration.

Example. Evaluate $\iint_G dx\,dy$ where G is the domain bounded by the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Here the integrand is identically equal to unity but nevertheless it is expedient to perform the change of variables

$$x = a\rho\cos\varphi, \qquad y = b\rho\sin\varphi \qquad (1.75)$$

The Jacobian of this transformation is equal to $ab\rho$. The domain of integration is changed under the mapping into the rectangle

$$0 \leqslant \varphi < 2\pi, \quad 0 \leqslant \rho \leqslant 1$$

Passing to the new variables and writing the double integral as an iterated one we obtain

$$\iint_G dx\, dy = ab \int_0^{2\pi} d\varphi \int_0^1 \rho\, d\rho = \pi ab$$

Exercises

1. Compute the area of the domain bounded by the curves $xy = 1$, $xy = 2$, $y = x^2$ and $y = 2x^2$.

Hint. Take, as new variables, the expressions

$$\xi = xy, \quad \eta = \frac{y}{x^2} \tag{1.76}$$

2. Draw the families of the coordinate curves corresponding to transformations (1.75) and (1.76).

7. Comparison with One-Dimensional Case. Integral Over an Oriented Domain. Formula (1.74) is analogous to the formula for changing variable in the definite integral:

$$\int_a^b f(x)\, dx = \int_\alpha^\beta f(x(t))\, x'(t)\, dt \tag{1.77}$$

The only difference between them is that in the case of one independent variable we do not take the absolute value of the derivative $x'(t)$ (which plays the role of the Jacobian here) but the derivative itself. This is accounted for by the fact that the definite integral $\int_a^b f(x)\, dx$ is taken over an oriented interval $[a, b]$ and changes its sign when the limits of integration are reversed whereas the domain of integration of a double integral is not oriented. If, for the definite integral, we introduce the condition that the limits of integration must be so set that the lower limit should be not greater than the upper, formula (1.77) (in the case of a monotone function $x = x(t)$ takes the form

$$\int_a^b f(x)\, dx = \int_\alpha^\beta f(x(t))\, |x'(t)|\, dt \tag{1.78}$$

(Check it up.)

On the other hand, for a double integral, we can also introduce the notion of an oriented domain and attach the sign plus or minus to its area according to the orientation.

We introduce the orientation of a domain by choosing a certain direction of describing the contour of its boundary as positive. Namely, the orientation of a domain is said to be *positive* if, when describing its boundary, the domain is always kept on the left of a person walking round the contour (see Fig. 1.32) and *negative* if otherwise. If the area of a (nonoriented) domain G is equal to S

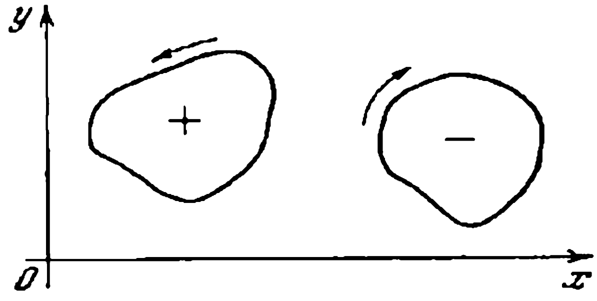


Fig. 1.32

we assume that the area of the oriented domain is equal to S if the orientation is positive and to $-S$ if negative. It can be shown that a mapping $x = x(\xi, \eta)$, $y = y(\xi, \eta)$ of a domain Γ on a domain G preserves the orientation if its Jacobian is positive and changes the orientation to the opposite if $\frac{D(x, y)}{D(\xi, \eta)} < 0$. Therefore the formula expressing the *area of the oriented domain* G in curvilinear coordinates is of the form

$$S = \iint\limits_{\Gamma} \frac{D(x, y)}{D(\xi, \eta)}\, d\xi\, d\eta$$

(i.e. the sign of modulus has been omitted), and formula (1.74) changes similarly.

2 Triple Integrals and Multiple Integrals of Higher Order

In the foregoing chapter we introduced the notion of a double integral. Here we are going to define the integral of a function of three independent variables, the so-called *triple integral*. Like double integrals, triple integrals are widely applied to various physical and geometrical problems. Some of the problems will be considered in § 3.

Triple integrals are in many respects almost completely analogous to double integrals and therefore we shall omit those proofs which do not essentially differ from the corresponding proofs of the theory of the double integral.

In § 5 of the present chapter we shall discuss the concept of *multiple integrals of higher order*, that is integrals of functions dependent on an arbitrary number of arguments.

§ 1. DEFINITION AND BASIC PROPERTIES OF TRIPLE INTEGRAL

1. Preliminary Observations. Volume of a Space Figure. The definitions of an interior point of a domain, a boundary, a closed domain, a diameter etc. given in § 1 of Chapter 1 for the plane are transferred without any changes to the case of the three-dimensional space.

When introducing the double integral we use the notion of area. Similarly, the definition of the triple integral is based on the notion of the volume of a space figure, a solid.

The reader is supposed to be familiar with the definition of the volume of a polyhedron known from elementary geometry. The extension of this notion to a wider class of figures can be performed in the same manner as it was done in § 1 of Chapter 1 where the notion of area was extended from polygonal figures to curvilinear squarable figures. Here we shall briefly present the corresponding arguments.

The volume $V(P)$ of a *polyhedral solid* (*polyhedral space figure*), i.e. a space figure composed of a finite number of polyhedrons, is a nonnegative quantity possessing the following properties:

1 (*monotonicity*). If P and Q are two polyhedral solids and P is contained in Q then

$$V(P) \leqslant V(Q)$$

2 (*additivity*). If P and Q are two polyhedral solids without interior points in common we have

$$V(P + Q) = V(P) + V(Q)$$

3 (*invariance*). If polyhedral solids P and Q are congruent their volumes are equal.

These three properties should be preserved when the concept of volume is extended from polyhedral solids to a wider class of space figures.

Take an arbitrary space figure* Φ and consider all the possible polyhedral figures embedded in it. The least upper bound of their volumes is referred to as the **interior (Jordan) content** of the figure Φ (if there is no nondegenerate polyhedral figure that can be embedded in the solid Φ we attribute, by definition, a zero interior content to Φ). Similarly, the greatest lower bound of the volumes of all polyhedral solids enveloping the figure Φ is called its **exterior (Jordan) content**. If the exterior and the interior contents of a figure Φ coincide their common value is said to be the **volume** of Φ. The following theorem is proved after a manner of Theorem 1.2:

Theorem 2.1. *For a space figure Φ to have volume it is necessary and sufficient that for every $\varepsilon > 0$ there exist two polyhedral figures $P \subset \Phi$ and $Q \supset \Phi$ such that*

$$V(Q) - V(P) < \varepsilon.$$

We say that a set is of **volume zero** if it can be embedded in a polyhedral solid of arbitrarily small volume. Using this notion we can rephrase theorem 2.1 as follows:

For a space figure Φ to have volume it is necessary and sufficient that its boundary have zero volume.

This criterion enables us to establish the existence of volume for some sufficiently wide classes of figures. For instance, the solids composed of a finite number of curvilinear cylinders having squarable lower bases and bounded above by surfaces defined by equations of the form $z = f(x, y)$, where $f(x, y)$ is a continuous function, form such a class. The volume of each cylinder is given by the double integral

$$\iint\limits_G f(x, y)\, dx\, dy$$

taken over the base of the cylinder.

* That is an arbitrary bounded set of points in space.

Another important class of such figures consists of space figures bounded by a finite number of smooth surfaces.* The proof of the fact that a solid bounded by smooth surfaces has volume is almost completely analogous to that of the fact that a smooth curve is of area zero, but since this proof involves some more complicated details we shall not present it here.

Repeating the argument given in § 1, Sec. 4 we can easily establish the validity of the following assertions:

(1) *Let* Φ_1 *and* Φ_2 *be two space figures having volume. Then their union* Φ *also has volume, and if the figures* Φ_1 *and* Φ_2 *have no interior points in common the volume of* Φ *is equal to the sum of the volumes of* Φ_1 *and* Φ_2.

(2) *The intersection (the meet) of two figures having volume also has volume.*

Note. We should pay attention to the fact that we deal with two different approaches to the concept of volume.

On the one hand, we have defined the volume of a curvilinear cylinder with a squarable base G bounded above by a surface $z = f(x, y)$ as the double integral

$$\iint_G f(x, y)\,dx\,dy$$

On the other hand, the concept of volume of a space figure has been introduced by approximating the figure with embedding and enveloping polyhedral figures. But it can be shown that these approaches are equivalent for a sufficiently wide class of figures, in particular, for the figures bounded by piecewise smooth surfaces.

2. Definition of Triple Integral. Let a bounded function $f(x, y, z)$ be defined on a space figure V which has volume.** Break up V into parts V_i, choose an arbitrary point (ξ_i, η_i, ζ_i) in each V_i and form the *integral sum*

$$T = \sum_{i=1}^{n} f(\xi_i, \eta_i, \zeta_i)\,\Delta v_i \tag{2.1}$$

where Δv_i is the volume of the element (subdomain) V_i and the sum is extended over all the elements of the partition. Let us introduce the following definitions.

*Definition 1. Let D be the maximal of the diameters $d(V_i)$ of the elements V_i into which the figure is divided (the **fineness** of the par-*

* A surface is said to be smooth if the tangent plane exists at each of its points and if the position of the tangent plane varies continuously as the point moves on the surface.

** In what follows we shall always suppose, without any further stipulation, that the space figures in question have volume.

tition). A number J is said to be the ***limit of integral sums*** (2.1), *as $D \to 0$, if for every $\varepsilon > 0$ there is $\delta > 0$ such that*

$$|T - J| < \varepsilon$$

if $D < \delta$.

In other words, the inequality $|T - J| < \varepsilon$ must hold for every integral sum T associated with any partition $\{V_i\}$ for which $D < \delta$ irrespective of the choice of the points (ξ_i, η_i, ζ_i) in each V_i.

Definition 2. *If the limit of integral sums* (2.1) *for $D \to 0$ exists it is called the* ***triple integral of the function*** *$f(x, y, z)$* ***over*** *V and denoted as*

$$\iiint\limits_V f(x, y, z)\,dv \quad \text{or} \quad \iiint\limits_V f(x, y, z)\,dx\,dy\,dz$$

In this case the function $f(x, y, z)$ is said to be **integrable** on V.

3. Conditions for Existence of Triple Integral. Integrability of Continuous Functions. As in the case of a function of one or two independent variables, an arbitrary bounded function $f(x, y, z)$ is by far not always integrable. To establish sufficient conditions for the existence of the triple integral we shall use upper and lower Darboux sums which are applied to single and double integrals.

Let $f(x, y, z)$ be a bounded function defined on a figure V and $\{V_i\}$ be a partition of the figure. Denote by M_i and m_i the respective upper and lower bounds of the values of the function $f(x, y, z)$ on V_i. The expressions

$$\sum_{i=1}^{n} M_i \Delta v_i \quad \text{and} \quad \sum_{i=1}^{n} m_i \Delta v_i$$

(where Δv_i is the volume of V_i) are called, respectively, the **upper** and the **lower Darboux sums** for the function $f(x, y, z)$ associated with the partition $\{V_i\}$ of the figure V. All the properties of the upper and the lower Darboux sums given in § 2 of Chapter 1 for two arguments are completely transferred to the case of three arguments.

The following necessary and sufficient condition for the existence of the triple integral is proved by applying arguments similar to those of the proof of Theorem 1.3:

Theorem 2.2. *A bounded function $f(x, y, z)$ defined on a space figure V is integrable on V if and only if for every $\varepsilon > 0$ there is a partition of the figure V such that the difference between the upper and the lower Darboux sums for the function $f(x, y, z)$ which correspond to the partition is less than ε.*

This criterion implies the following theorems similar to Theorems 1.4 and 1.4′ proved for double integrals.

Theorem 2.3. *Each continuous function $f(x, y, z)$ defined in a bounded closed domain* V is integrable over the domain.*

Theorem 2.4. *If a function $f(x, y, z)$ is bounded throughout a bounded closed domain, possibly except the points belonging to a set of volume zero, the function $f(x, y, z)$ is integrable on the domain.*

4. Properties of Triple Integral. The basic properties of the triple integral are completely analogous to those of the double integral. We now enumerate them.

1-2 (*linearity*). If $f_1(x, y, z)$ and $f_2(x, y, z)$ are integrable over a domain V and k_1 and k_2 are constants the function $k_1f_1 + k_2f_2$ is also integrable on V and we have

$$\iiint\limits_V \left[k_1f_1(x, y, z) + k_2f_2(x, y, z)\right]dv =$$

$$= k_1 \iiint\limits_V f_1(x, y, z)\,dv + k_2 \iiint\limits_V f_2(x, y, z)\,dv$$

3 (*additivity*). If V is the union of two space figures V_1 and V_2 with no interior points in common and a function $f(x, y, z)$ is integrable over V_1 and V_2, then $f(x, y, z)$ is integrable over V and

$$\iiint\limits_V f(x, y, z)\,dv = \iiint\limits_{V_1} f(x, y, z)\,dv + \iiint\limits_{V_2} f(x, y, z)\,dv$$

4 (*monotonicity*). If $f_1(x, y, z) \geqslant f_2(x, y, z)$ and both functions are integrable on V the inequality

$$\iiint\limits_V f_1(x, y, z)\,dv \geqslant \iiint\limits_V f_2(x, y, z)\,dv$$

takes place.

5 (*estimation of the modulus of the integral*). If $f(x, y, z)$ is integrable on V the function $|f(x, y, z)|$ is also integrable and

$$\left|\iiint\limits_V f(x, y, z)\,dv\right| \leqslant \iiint\limits_V |f(x, y, z)|\,dv$$

6 (*the mean value theorem*). If a function $f(x, y, z)$ is integrable on V and satisfies the inequality $m \leqslant f(x, y, z) \leqslant M$ then

$$mv \leqslant \iiint\limits_V f(x, y, z)\,dv \leqslant Mv$$

where v is the volume of V.

* Here and henceforward the term "domain" is used as a synonym for a space figure which has volume.

For the case of a continuous function the mean value theorem can be rephrased as follows:

6'. If a function $f(x, y, z)$ is continuous and V is a bounded closed connected domain, there is a point (ξ, η, ζ) in the domain V such that

$$\iiint_V f(x, y, z)\, dv = f(\xi, \eta, \zeta)\, v$$

5. Triple Integral as an Additive Set Function. By analogy with a set function defined for plane figures, we can introduce the notion of a set function whose argument is a space figure (domain). An example of such a function (defined for each space figure that has volume) is the volume of a domain. If the whole space or its part is occupied by a substance we can associate with each domain the mass contained in it and thus obtain a set function defined for domains in space. Volume and mass possess the property of additivity which is formulated in the same manner as in the case of plane figures: a set function $F(V)$ is said to be *additive* if, for any two domains V_1 and V_2 having no interior points in common, on which $F(V)$ is defined, the value $F(V_1 + V_2)$ is also defined and

$$F(V_1 + V_2) = F(V_1) + F(V_2)$$

If $f(x, y, z)$ is an integrable function the triple integral

$$\iiint_V f(x, y, z)\, dv$$

regarded as a function of its domain of integration is an additive set function (see property 3 in Sec. 4).

The notion of the derivative of an additive set function defined in space with respect to volume is introduced by analogy with the two-dimensional case. Namely, a number A is said to be the limit of the quotient

$$\frac{F(V)}{v}$$

(where v is the volume of V), as V is contracted to a point M_0, if for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left|\frac{F(V)}{v} - A\right| < \varepsilon$$

for every domain V entirely contained within the sphere of radius δ with centre at M_0. The limit is called the *derivative of the function $F(V)$ with respect to volume at the point M_0* and denoted by the symbol

$$\lim_{V \to M_0} \frac{F(V)}{v} \quad \text{or} \quad \frac{dF}{dv}$$

If $F(V)$ is the mass contained in the domain V its derivative with respect to volume (if it exists) is the volume density $\rho(x, y, z)$ of mass distribution in space.

From the mean value theorem for the triple integral it follows that if the integrand is continuous the derivative of the integral with respect to volume exists and coincides with the integrand:

$$\frac{d}{dv}\iiint\limits_V f(x, y, z)\, dv = f(x, y, z)$$

In this case the integral is the only additive set function in space whose derivative with respect to volume is equal to the continuous function $f(x, y, z)$.

§ 2. SOME APPLICATIONS OF TRIPLE INTEGRAL IN PHYSICS AND GEOMETRY

We now consider some typical problems involving computation of triple integrals.

1. Computing Volumes. If a space figure V has volume the triple integral

$$\iiint\limits_V dx\, dy\, dz \tag{2.2}$$

is equal to the volume. Indeed, each integral sum corresponding to the integral equals the volume. Triple integrals are sometimes more convenient than double integrals, for computing volumes, because they enable us to put down the expression of the volume not only for a curvilinear cylinder but for an arbitrary solid as well.

2. Finding the Mass of a Solid from Its Density. If we are given a solid with a volume density $\rho(x, y, z)$ of mass distribution which is a continuous function the triple integral

$$\iiint\limits_V \rho(x, y, z)\, dx\, dy\, dz$$

taken over the entire volume occupied by the solid gives the mass of the solid. The derivation of this formula is completely analogous to that of the formula for determining the mass of a plate from its surface density.

3. Moment of Inertia. Performing the usual passage to the limit from a system of mass points to a continuously distributed mass we can easily derive the following expressions for the moments of inertia of a solid with volume density $\rho(x, y, z)$ about the coordi-

nate axes:

$$I_x = \iiint\limits_V (y^2 + z^2)\,\rho(x, y, z)\,dx\,dy\,dz$$

$$I_y = \iiint\limits_V (x^2 + z^2)\,\rho(x, y, z)\,dx\,dy\,dz$$

$$I_z = \iiint\limits_V (x^2 + y^2)\,\rho(x, y, z)\,dx\,dy\,dz$$

The moment of inertia about the origin of coordinates is given by the formula

$$I_0 = \iiint\limits_V (x^2 + y^2 + z^2)\,\rho(x, y, z)\,dx\,dy\,dz$$

4. **Determining the Coordinates of the Centre of Gravity.** The coordinates of the centre of gravity of a solid with mass density $\rho(x, y, z)$ are expressed by the formulas

$$x_c = \frac{\iiint\limits_V x\rho(x, y, z)\,dx\,dy\,dz}{\iiint\limits_V \rho(x, y, z)\,dx\,dy\,dz}, \quad y_c = \frac{\iiint\limits_V y\rho(x, y, z)\,dx\,dy\,dz}{\iiint\limits_V \rho(x, y, z)\,dx\,dy\,dz},$$

$$z_c = \frac{\iiint\limits_V z\rho(x, y, z)\,dx\,dy\,dz}{\iiint\limits_V \rho(x, y, z)\,dx\,dy\,dz}$$

which are received by means of the same arguments as in the case of two dimensions. In particular, if the solid is homogeneous, i.e. $\rho(x, y, z) = \text{const}$, the formulas for the coordinates of the centre of gravity are simplified:

$$x_c = \frac{\iiint\limits_V x\,dv}{\iiint\limits_V dv}, \quad y_c = \frac{\iiint\limits_V y\,dv}{\iiint\limits_V dv}, \quad z_c = \frac{\iiint\limits_V z\,dv}{\iiint\limits_V dv}$$

5. **Gravitational Attraction of a Material Point by a Solid.** Suppose we are given a solid occupying a domain V and having a density $\rho(x, y, z)$ and a material point (lying outside V) of mass m with the coordinates (x_0, y_0, z_0). Let us determine the gravitational force with which the material point is attracted by the solid. Consider an element of volume dv of the solid. The mass of the element is equal to $\rho(x, y, z)\,dv$, and it attracts the material point with a force whose numerical value is equal to

$$\gamma \frac{m\rho(x, y, z)\,dv}{r^2}$$

where γ is the *gravitational constant* (whose value depends on the choice of the system of units) and

$$r = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$$

the direction of the force coinciding with the direction of the vector $\mathbf{r}$ joining the points (x_0, y_0, z_0) and (x, y, z). Consider the projection of the force on the x-axis. It is equal to

$$\gamma \frac{(x-x_0)\, m\rho(x, y, z)\, dv}{r^3} \tag{2.3}$$

$\left(\text{because the cosine of the angle between the axis and the vector } \mathbf{r} \text{ is equal to } \frac{x-x_0}{r}\right)$. To evaluate the projection F_x on the x-axis of the force with which the entire solid attracts the material point we must sum up elements (2.3), i.e. compute the corresponding triple integral.

Thus, we have

$$F_x = \iiint\limits_V \frac{\gamma m\rho(x, y, z)(x-x_0)}{r^3}\, dv$$

The other two projections are found similarly:

$$F_y = \iiint\limits_V \frac{\gamma m\rho(x, y, z)(y-y_0)}{r^3}\, dv$$

$$F_z = \iiint\limits_V \frac{\gamma m\rho(x, y, z)(z-z_0)}{r^3}\, dv$$

Note. It should be taken into account that, from the mathematical point of view, the formulas obtained here and in § 4 of the preceding chapter for similar problems are in fact the definitions of the corresponding notions (centre of gravity, moment of inertia etc.) for the case of a continuously distributed mass. The justification of these definitions does not lie in logical arguments but is based on the fact that the results of the corresponding experiments are coincident with those obtained by calculations performed according to the formulas.

§ 3. EVALUATING TRIPLE INTEGRAL

As in the case of the double integral, the main technique used for evaluating a triple integral is based on reducing the integral to an iterated (repeated) one, i.e. on replacing the integration over a space figure by successive separate integrations with respect to each variable.*

---

* Here we mean the exact analytic computation of an integral. In practical approximate calculations the reduction of a multiple integral to an iterated one is rarely applied.

We shall first consider a special case of the problem of reducing a triple integral to an iterated one when the domain of integration

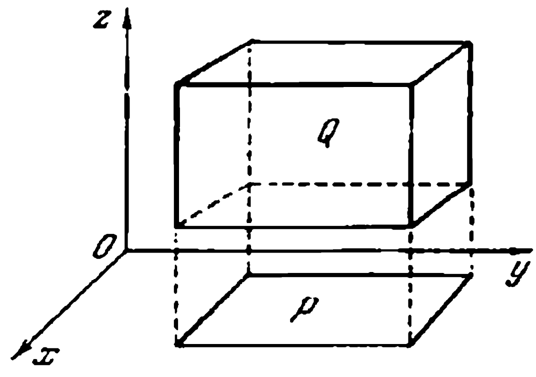


Fig. 2.1

is a rectangular parallelepiped with faces parallel to the coordinate planes of a Cartesian coordinate system.

1. Reducing Triple Integral Over a Rectangular Parallelepiped to an Iterated Integral. Consider a triple integral

$$\iiint\limits_{Q} f(x, y, z)\,dx\,dy\,dz$$

whose domain of integration Q is a rectangular parallelepiped determined by inequalities

$$a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d, \quad k \leqslant z \leqslant l$$

(see Fig. 2.1), the projection of the parallelepiped on the x, y-plane being the rectangle P specified by the relations

$$a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d$$

The following theorem takes place:

Theorem 2.5. If, for a function $f(x, y, z)$, the triple integral

$$\iiint\limits_{Q} f(x, y, z)\,dv$$

exists and if the integral

$$I(x, y) = \int\limits_{k}^{l} f(x, y, z)\,dz$$

exists for every fixed point (x, y) of P, the iterated integral

$$\iint\limits_{P} dx\,dy \int\limits_{k}^{l} f(x, y, z)\,dz$$

also exists and there is a relation of the form

$$\iiint\limits_{Q} f(x, y, z)\,dv = \iint\limits_{P} dx\,dy \int\limits_{h}^{l} f(x, y, z)\,dz \tag{2.4}$$

The proof of the theorem is similar to that of the theorem on the reduction of a double integral to an iterated one (see Theorem 1.5). It is based on establishing the fact that every integral sum corresponding to the integral $\iint\limits_{P} I(x, y)\,dx\,dy$ for an arbitrary partition lies between the upper and the lower Darboux sums associated with the triple integral $\iiint\limits_{Q} f(x, y, z)\,dv$.

If we assume that the integral

$$J(x) = \int\limits_{c}^{d} I(x, y)\,dy$$

also exists for each fixed x from the interval $a \leqslant x \leqslant b$ we can substitute, in formula (2.4), the successive integrations performed first with respect to y and then to x for the integration over the rectangle P. This enables us to rewrite equality (2.4) in the form

$$\iiint\limits_{Q} f(x, y, z)\,dv = \int\limits_{a}^{b} dx \int\limits_{c}^{d} dy \int\limits_{h}^{l} f(x, y, z)\,dz \tag{2.5}$$

It is this formula that reduces the evaluation of a triple integral over a parallelepiped Q to successive separate integrations with respect to each variable. In the expression on the right-hand side of formula (2.5) the first integration is performed with respect to z, the second with respect to y and, finally, the third with respect to x. If we suppose that the integrals

$$I_1(y, z) = \int\limits_{a}^{b} f(x, y, z)\,dx \quad \text{and} \quad J_1(z) = \int\limits_{c}^{d} I_1(y, z)\,dy$$

exist we can derive the analogous formula

$$\iiint\limits_{Q} f(x, y, z)\,dv = \int\limits_{h}^{l} dz \int\limits_{c}^{d} dy \int\limits_{a}^{b} f(x, y, z)\,dx$$

Similarly, on condition that the corresponding single and double integrals exist, we can establish analogous formulas reducing the triple integral to an iterated one taken with respect to x, y and z in various orders. In particular, if the function $f(x, y, z)$ is conti

nuous the triple integral and all the possible double and single integrals are sure to exist and therefore when evaluating a triple integral of a continuous function we can separately integrate with respect to the variables x, y and z in any orders and combinations. In the general case of an arbitrary integrable function the orders are not always interchangeable.

2. **Reducing Triple Integral Over a Curvilinear Domain to an Iterated Integral.** We now consider a curvilinear domain V bounded above and below by the surfaces

$$z = z_1(x, y) \quad \text{and} \quad z = z_2(x, y)$$

$(z_2(x, y) \geqslant z_1(x, y))$ and on the sides by a cylindrical surface. Denote by G the projection of the domain V on the x, y-plane (Fig. 2.2). A feature of such a domain is that each straight line parallel to the

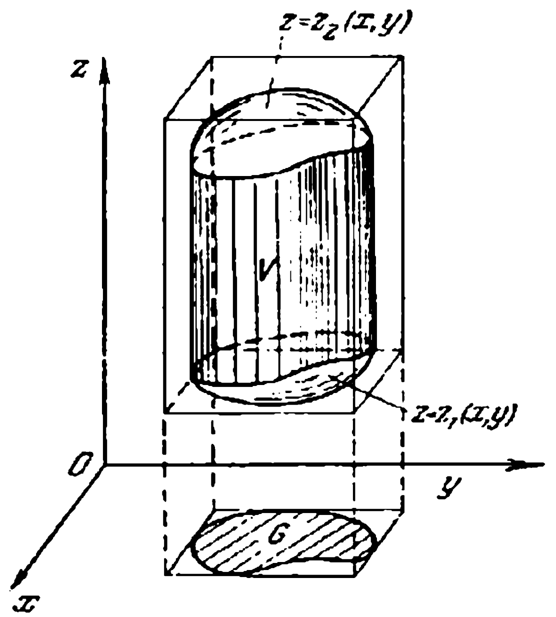


Fig. 2.2

z-axis and passing through a point of the domain has at most two common points with its boundary. For brevity, we shall call such a domain *regular in the z-direction*. Let a function $f(x, y, z)$ defined in the domain V be integrable. Suppose that for every fixed point (x, y) of G the single integral

$$\int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\, dz$$

exists. Let us embed the domain V in a rectangular parallelepiped Q determined by inequalities

$$a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d, \quad k \leqslant z \leqslant l$$

and define in Q an auxiliary function $f^*(x, y, z)$ by putting

$$f^*(x, y, z) = \begin{cases} f(x, y, z) & \text{in } V \\ 0 & \text{outside } V \end{cases}$$

The function $f^*(x, y, z)$ is obviously integrable on Q and

$$\iiint\limits_{Q} f^*(x, y, z)\,dv = \iiint\limits_{V} f(x, y, z)\,dv \tag{2.6}$$

Applying formula (2.4) to $f^*(x, y, z)$ we obtain

$$\iiint\limits_{Q} f^*(x, y, z)\,dv = \iint\limits_{P} dx\,dy \int\limits_{k}^{l} f^*(x, y, z)\,dz \tag{2.7}$$

where P is the projection of Q on the x, y-plane.

The function $f^*(x, y, z)$ vanishing outside V, we thus have

$$\int\limits_{k}^{l} f^*(x, y, z)\,dz = \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz \tag{2.8}$$

Expression (2.8) is a function of x and y which is identically equal to zero outside the domain G. Therefore the double integral of the expression taken over P (the projection of the parallelepiped Q on the x, y-plane) coincides with the integral of the expression taken over G. Therefore, taking into account formulas (2.6) and (2.8) we can rewrite equality (2.7) in the form

$$\iiint\limits_{V} f(x, y, z)\,dx\,dy\,dz = \iint\limits_{G} dx\,dy \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz \tag{2.9}$$

Hence, we have arrived at the following result:

Theorem 2.6. If the triple integral

$$\iiint\limits_{V} f(x, y, z)\,dv$$

exists for a function $f(x, y, z)$ defined in a domain V regular in the z-direction and if the integral

$$I(x, y) = \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz$$

exists for each fixed point (x, y) belonging to the projection G of the domain V on the x, y-plane, the iterated integral

$$\iint\limits_{G} dx\,dy \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz$$

also exists and equality (2.9) holds.

The expression

$$I(x, y) = \int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\, dz$$

is a function of two variables. If, for the function $f(x, y, z)$ and for the domain G over which it is integrated, the conditions of Theorem 1.6 are fulfilled, the double integral

$$\iint_G I(x, y)\, dx\, dy$$

can be represented as an iterated integral taken, for instance, first with respect to y and then with respect to x. This results in the relation

$$\iiint_V f(x, y, z)\, dv = \int_a^b dx \int_{y_1(x)}^{y_2(x)} dy \int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\, dz \qquad (2.10)$$

It is this final formula that reduces the triple integral to an iterated one. The variables x, y and z can be interchanged if the corresponding conditions hold. For example, we can reduce the triple integral to an iterated one in which the integration is performed in a different order, for instance, first with respect to x, then with respect to y and, finally, with respect to z. In all the possible cases the limits of integration with respect to a variable are dependent on those variables with respect to which the integration has not yet been performed.

When deriving formulas (2.9) we have taken advantage of the fact that each straight line parallel to the z-axis and passing through a point of the domain G cuts its boundary at no more than two points. If the domain is of a more complicated form then, to reduce a triple integral taken over it to an iterated integral, we must break up the domain into parts such that to each of them formula (2.9) is applicable. We have already encountered such a situation in the case of a double integral.

Now, summing up, we can briefly formulate the rule of reducing a triple integral to an iterated one (for definiteness, we suppose that the iterated integral is first taken with respect to z and then with respect to the other variables).

1. Break the domain over which the triple integral is taken into subdomains such that every vertical (parallel to the z-axis) straight line passing through a point of a subdomain has at most two common points with its boundary. In what follows we mention only one such subdomain.

2. Fix arbitrary x and y and consider the corresponding straight line parallel to the z-axis. Let $z_1(x, y)$ and $z_2(x, y)$ be the z-coordina-

tes of the points of intersection of the straight line with the boundary of the domain (subdomain) of integration. The expressions $z_1(x, y)$ and $z_2(x, y)$ should be taken as the limits of integration with respect to z.

3. Take the function of two variables x and y obtained after the integration with respect to z has been performed. The domain of definition of the function is the projection of the space figure V (or of the corresponding subdomain of V) on the x, y-plane. Finally, replace the double integral of this function of two variables by the corresponding iterated integral following the rules described in § 5 of Chapter 1.

The formula for reducing a triple integral to an iterated one is essentially based on the process of grouping the summands which has already been dealt with. Instead of summing up the elements of integration $f(x, y, z)\,dx\,dy\,dz$ in an arbitrary fashion $\left(\text{i.e. evaluating the integral } \iiint\limits_V f(x, y, z)\,dx\,dy\,dz\right)$ we first collect all the summands corresponding to one elementary cylinder with base in the vicinity of a point (x, y) (that is, we take the integral $\int\limits_{z_1(x,y)}^{z_2(x,y)} f(x, y, z)dx$), then we add together the results corresponding to a section of the domain V by a plane $x = \text{const}$ (which means that we compute the integral $\int\limits_{y_1(x)}^{y_2(x)} dy \int\limits_{z_1(x,y)}^{z_2(x,y)} f(x, y, z)\,dz$) and, finally, add up all the quantities thus obtained that correspond to all the sections which results in the formula

$$\iiint\limits_V f(x, y, z)\,dv = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} dy \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz$$

Exercise. Write the triple integral of a function $f(x, y, z)$ taken over the sphere (ball)

$$x^2 + y^2 + z^2 \leqslant a^2$$

in the form of an iterated integral.

Answer.

$$\iiint\limits_{x^2+y^2+z^2 \leqslant a^2} f(x, y, z)\,dv = \int\limits_{-a}^{a} dx \int\limits_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} dy \int\limits_{-\sqrt{a^2-x^2-y^2}}^{\sqrt{a^2-x^2-y^2}} f(x, y, z)\,dz$$

§ 4. CHANGE OF VARIABLES IN TRIPLE INTEGRAL

We have already dealt with the method of changing variables for the double integral (see § 6 of Chapter 1). For the change of

variable in a single integral the reader can be referred to [8], Chapter 6, § 2. Here we shall consider the problem of changing variables in the triple integral. The subject matter we are going to discuss is in many respects similar to that of § 6, Chapter 1.

1. Mapping of Space Figures. Consider two specimens of a three-dimensional space. Introduce a Cartesian coordinate system x, y, z in one of them and ξ, η, ζ in the other. Furthermore, let V and Ω be domains belonging to the spaces, their boundaries being, respectively, some piecewise smooth surfaces S and Σ (Fig. 2.3). Suppose

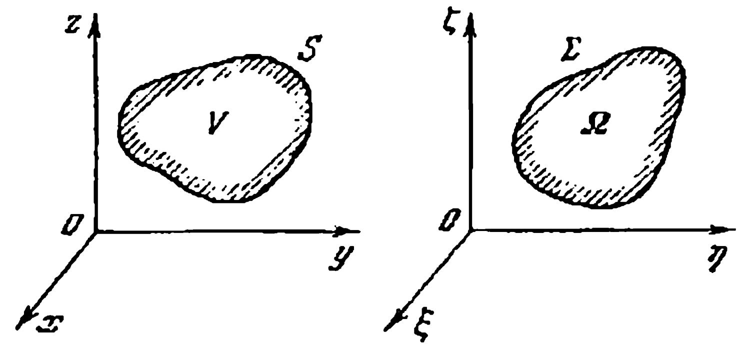


Fig. 2.3

that there is a correspondence between the points of the domains which is one-to-one and continuous in both directions. The correspondence can be expressed by means of three functions

$$x = x(\xi, \eta, \zeta), \quad y = y(\xi, \eta, \zeta), \quad z = z(\xi, \eta, \zeta) \qquad (2.11)$$

or by the inverse functions

$$\xi = \xi(x, y, z), \quad \eta = \eta(x, y, z), \quad \zeta = \zeta(x, y, z) \qquad (2.12)$$

We shall consider functions (2.11) and (2.12) not only to be continuous but also possessing the continuous first-order partial derivatives. Then the Jacobians

$$\frac{D(x, y, z)}{D(\xi, \eta, \zeta)} \quad \text{and} \quad \frac{D(\xi, \eta, \zeta)}{D(x, y, z)}$$

exist and are continuous. Let each Jacobian be different from zero. These conditions imply the relation

$$\frac{D(x, y, z)}{D(\xi, \eta, \zeta)} \cdot \frac{D(\xi, \eta, \zeta)}{D(x, y, z)} = 1 \qquad (2.13)$$

As in the two-dimensional case, we can show that under the mapping specified by the correspondence which is determined by formulas (2.11) and (2.12) the interior points of one domain go into the interior points of the other and the boundary points into the boundary ones.

2. Curvilinear Coordinates in Space. Mapping (2.11) transforms the domain Ω into V. Consequently, the specification of a point

(ξ, η, ζ) belonging to Ω uniquely determines the corresponding point (x, y, z) of V. In other words, the quantities ξ, η and ζ can be regarded as coordinates (different from Cartesian ones, in the general case) of the points of the domain V. They are called **curvilinear coordinates.**

Consider, in the domain Ω, a plane determined by the relation $\xi = \xi_0$, i.e. a plane parallel to the coordinate plane η, ζ. Under mapping (2.11), the plane goes into a surface lying in the domain V. The Cartesian coordinates of the points lying on the surface are expressed by the formulas

$$x = x(\xi_0, \eta, \zeta), \quad y = y(\xi_0, \eta, \zeta), \quad z = z(\xi_0, \eta, \zeta)^* \qquad (2.14)$$

Making ξ_0 assume all the possible values we obtain a one-parameter family of surfaces, the parameter being ξ. The planes η = const and ζ = const are similarly mapped onto two families of surfaces lying in the domain V. These three families form the set of the so-called

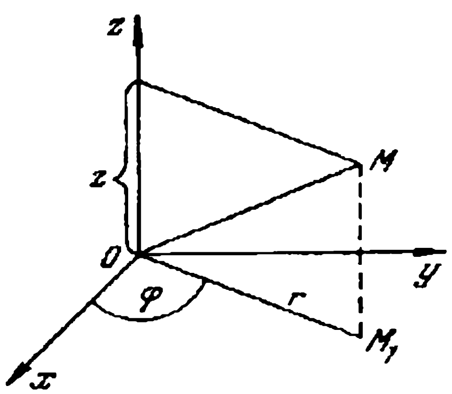


Fig. 2.4

coordinate surfaces. On condition that mapping (2.11) is one-to-one only a single surface belonging to one of the families passes through every given point of the domain V.

3. Cylindrical and Spherical Coordinates. We shall consider two curvilinear coordinate systems in space which are used most frequently, namely, cylindrical and spherical coordinates.

(a) *Cylindrical coordinates.* Let us specify the position of an arbitrary point M in space by means of its Cartesian coordinate z and polar coordinates r, φ of its projection M_1 on the x, y-plane (see Fig. 2.4). The quantities r, φ, z are referred to as the **cylindrical coordinates** of the point M. The figure directly implies that they are connected with the Cartesian coordinates of the point M by the relationship

$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad z = z \qquad (2.15)$$

---

* Expressions (2.14) are the so-called parametric equations of the surface. For a more detailed discussion of parametric equations of a surface see Chapter 3.

We have the following three families of coordinate surfaces corresponding to the cylindrical coordinates:

(α) the cylinders $r = \text{const}$ $(0 \leqslant r < \infty)$.

(β) the vertical half-planes $\varphi = \text{const}$ $(0 \leqslant \varphi < 2\pi)$,

(γ) the horizontal planes $z = \text{const}$ $(-\infty < z < \infty)$.

The Jacobian of the transformation from the Cartesian coordinates to the cylindrical ones is equal to

$$\frac{D(x, y, z)}{D(r, \varphi, z)} = \begin{vmatrix} \cos\varphi & \sin\varphi & 0 \\ -r\sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r \tag{2.16}$$

Formulas (2.15) expressing the relationship between Cartesian and cylindrical coordinates determine a mapping of the domain

$$0 \leqslant r < \infty, \quad 0 \leqslant \varphi < 2\pi, \quad -\infty < z < \infty \tag{2.17}$$

lying in the r, φ, z-space onto the entire x, y, z-space. Under the mapping, to each point $(0, 0, z_0)$ there corresponds an entire half-segment of the form

$$r = 0, \quad 0 \leqslant \varphi < 2\pi, \quad z = z_0$$

belonging to domain (2.17). Therefore the mapping is not one-to-one at the points lying on the z-axis. But it is obviously one-to-one at all the other points of the x, y, z-space.

(b) *Spherical coordinates.* Let the position of the moving point M in space be determined by the following three quantities:

(α) the distance ρ from the origin O to M.

(β) the angle θ between the line segment OM and the direction of the positive half-axis z,

(γ) the angle φ formed by the projection OM_1 of the line segment OM on the x, y-plane and the positive direction of the x-axis

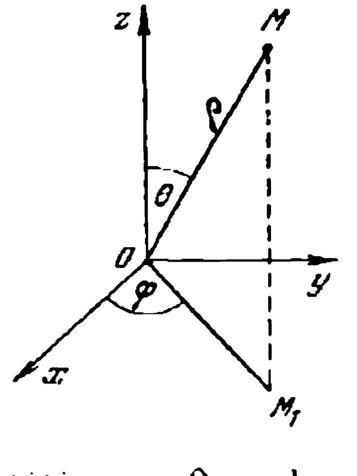


Fig. 2.5

(see Fig. 2.5). The quantities ρ, θ and φ are called the spherical coordinates of the point M. The figure shows that the relationship between the Cartesian coordinates of the point M and its spherical

coordinates is expressed by the formulas

$$x = \rho \sin \theta \cos \varphi, \quad y = \rho \sin \theta \sin \varphi, \quad z = \rho \cos \theta \qquad (2.18)$$

The three families of the coordinate surfaces corresponding to spherical coordinates are:

(α) the spheres $\rho = \text{const}$ $(0 \leqslant \rho < \infty)$,
(β) the conical nappes (semi-cones) $\theta = \text{const}$ $(0 \leqslant \theta \leqslant \pi)$,
(γ) the vertical half-planes $\varphi = \text{const}$ $(0 \leqslant \varphi < 2\pi)$.

The Jacobian of the transformation of Cartesian coordinates into spherical coordinates is equal to

$$\frac{D(x, y, z)}{D(\rho, \varphi, \theta)} = \begin{vmatrix} \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \\ \rho\cos\theta\cos\varphi & \rho\cos\theta\sin\varphi & -\rho\sin\theta \\ -\rho\sin\theta\sin\varphi & \rho\sin\theta\cos\varphi & 0 \end{vmatrix} = \rho^2 \sin\theta \qquad (2.19)$$

Formulas (2.18) determine a mapping of the domain

$$0 \leqslant \rho < \infty, \quad 0 \leqslant \theta \leqslant \pi, \quad 0 \leqslant \varphi < 2\pi$$

("a semi-infinite rod") of the ρ, θ, φ-space onto the entire x, y, z-space. Like the mapping corresponding to cylindrical coordinates, this mapping is one-to-one everywhere except the points lying on the z-axis. Indeed, each point $(0, 0, z_0)$ is the image, under the mapping, of the half-segment $\rho = z_0$, $\theta = 0$, $0 \leqslant \varphi < 2\pi$ for $z_0 > 0$, and of the half-segment $\rho = z_0$, $\theta = \pi$, $0 \leqslant \varphi < 2\pi$ for $z_0 < 0$, and the entire rectangle $\rho = 0$, $0 \leqslant \theta \leqslant \pi$, $0 \leqslant \varphi < 2\pi$ is mapped into the point $(0, 0, 0)$.

4. Element of Volume in Curvilinear Coordinates. Let us find the expression of an element of volume in curvilinear coordinates. Consider again a space figure V in which some curvilinear coordinates ξ, η, ζ are introduced, and let the formulas

$$x = x(\xi, \eta, \zeta), \quad y = y(\xi, \eta, \zeta), \quad z = z(\xi, \eta, \zeta) \qquad (2.20)$$

express the relationship between ξ, η, ζ and the Cartesian coordinates x, y, z. The functions $x(\xi, \eta, \zeta)$, $y(\xi, \eta, \zeta)$ and $z(\xi, \eta, \zeta)$ are supposed to be continuous and possess continuous derivatives. We also impose the condition that the Jacobian should be different from zero.

Consider three pairs of coordinate surfaces drawn infinitely close to each other. Let the first pair be given by fixed values of the first coordinate which are equal, respectively, to ξ and $\xi + d\xi$. Similarly, let the second pair be specified by values η and $\eta + d\eta$ of the second coordinate, and the third one by values ζ and $\zeta + d\zeta$ of the third coordinate. These three pairs of surfaces cut out of space an infinitesimal curvilinear parallelepiped. We shall evaluate its volume dv neglecting the terms of the second and higher order of

smallness relative to the volume. This parallelepiped coincides, to within infinitesimals of order higher than the first, with the rectilinear parallelepiped whose edges are the vectors $\overline{PP}_1$, $\overline{PP}_2$ and $\overline{PP}_3$ (Fig. 2.6). It can easily be shown that the vectors have the following coordinates:

$$\overline{PP_1} = \left(\frac{\partial x}{\partial \xi} d\xi, \quad \frac{\partial y}{\partial \xi} d\xi, \quad \frac{\partial z}{\partial \xi} d\xi\right),$$

$$\overline{PP_2} = \left(\frac{\partial x}{\partial \eta} d\eta, \quad \frac{\partial y}{\partial \eta} d\eta, \quad \frac{\partial z}{\partial \eta} d\eta\right),$$

$$\overline{PP_3} = \left(\frac{\partial x}{\partial \zeta} d\zeta, \quad \frac{\partial y}{\partial \zeta} d\zeta, \quad \frac{\partial z}{\partial \zeta} d\zeta\right)$$

where we have again restricted ourselves to the principal terms (i.e. of the first order of smallness). As is well known, the volume

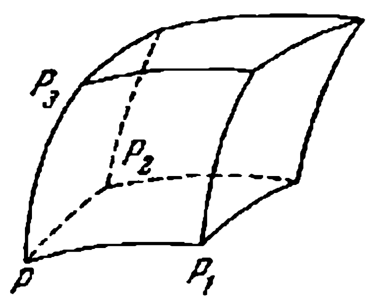


Fig. 2.6

of a parallelepiped constructed on three vectors is equal to the absolute value of the determinant with the coordinates of the vectors as its elements. Consequently, we have

$$dv = \pm \begin{vmatrix} \frac{\partial x}{\partial \xi} d\xi & \frac{\partial y}{\partial \xi} d\xi & \frac{\partial z}{\partial \xi} d\xi \\ \frac{\partial x}{\partial \eta} d\eta & \frac{\partial y}{\partial \eta} d\eta & \frac{\partial z}{\partial \eta} d\eta \\ \frac{\partial x}{\partial \zeta} d\zeta & \frac{\partial y}{\partial \zeta} d\zeta & \frac{\partial z}{\partial \zeta} d\zeta \end{vmatrix} = \pm \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} & \frac{\partial z}{\partial \eta} \\ \frac{\partial x}{\partial \zeta} & \frac{\partial y}{\partial \zeta} & \frac{\partial z}{\partial \zeta} \end{vmatrix} d\xi \, d\eta \, d\zeta$$

where the sign + or — is so chosen that the whole expression is positive. Thus, we see that

$$dv = | J(\xi, \eta, \zeta) | \, d\xi \, d\eta \, d\zeta \tag{2.21}$$

where $J(\xi, \eta, \zeta) = \frac{D(x, y, z)}{D(\xi, \eta, \zeta)}$ is the Jacobian of transformation (2.20).

5. Change of Variables in Triple Integral. Geometric Meaning of the Jacobian. We have shown that the volume of an infinitesimal element is expressed in curvilinear coordinates by formula (2.21). It follows directly that the volume of a finite domain V can be

written as the triple integral

$$\iiint_{\Omega} |J(\xi, \eta, \zeta)|\, d\xi\, d\eta\, d\zeta \qquad (2.22)$$

taken over the range Ω, of the variables ξ, η and ζ, which is mapped onto the domain V under mapping (2.20).*

On the basis of the expression for the volume we can derive the formula of changing variables by means of the arguments given below which are similar to the ones presented in § 6, Sec. 6 of Chapter 1.

Let $f(x, y, z)$ be a continuous function defined in a bounded closed domain V. Under these conditions, the integral

$$\iiint_{V} f(x, y, z)\, dx\, dy\, dz \qquad (2.23)$$

exists and is equal to the limit of integral sums of the form

$$\sum_{i=1}^{n} f(x_i, y_i, z_i)\, \Delta v_i \qquad (2.24)$$

Suppose formula (2.20) establishes a correspondence between the points of the domain V and a domain Ω which is the range of the variables ξ, η and ζ. Besides, let the correspondence satisfy the conditions enumerated in Sec. 1. This correspondence attributes to each partition $\{V_i\}$ of the domain V into parts V_i a certain partition $\{\Omega_i\}$ of the domain Ω and vice versa. According to (2.22), the volume Δv_i of the subdomain V_i is representable in the form

$$\Delta v_i = \iiint_{\Omega_i} |J(\xi, \eta, \zeta)|\, d\xi\, d\eta\, d\zeta$$

Applying the mean value theorem to the integral we receive

$$\Delta v_i = |J(\xi_i^*, \eta_i^*, \zeta_i^*)|\, \Delta\omega_i$$

where $\Delta\omega_i$ is the volume of the subdomain Ω_i and $(\xi_i^*, \eta_i^*, \zeta_i^*)$ is a point belonging to Ω_i.

Each point (x_i, y_i, z_i) entering into sum (2.24) can be chosen quite arbitrarily within the corresponding subdomain V_i. In parti-

* Here we have left out the calculations which are similar to those written in full in § 6 of Chapter 1 for the case of two independent variables.

If the reader carefully studies Theorem 1.7 it will not be difficult for him to prove formula (2.22). Here, as in the case of dimension 2, the basic idea lies in approximating a nonlinear mapping of a small domain by an appropriate linear mapping.

cular, we can take, as (x_i, y_i, z_i), the point whose curvilinear coordinates are ξ_i^*, η_i^* and ζ_i^*. Therefore integral sum (2.24) can be rewritten in the form

$$\sum_{i=1}^{n} f(x(\xi_i^*, \eta_i^*, \zeta_i^*), \quad y(\xi_i^*, \eta_i^*, \zeta_i^*), \quad z(\xi_i^*, \eta_i^*, \zeta_i^*))\,|J(\xi_i^*, \eta_i^*, \zeta_i^*)|\,\Delta\omega_i \tag{2.25}$$

that is as an integral sum corresponding to the integral

$$\iiint_{\Omega} f(x(\xi, \eta, \zeta), \quad y(\xi, \eta, \zeta), \quad z(\xi, \eta, \zeta))\,|J(\xi, \eta, \zeta)|\,d\xi\,d\eta\,d\zeta \tag{2.26}$$

The last integral is sure to exist because its integrand is continuous. Consider a sequence of partitions $\{V_i\}$ of the domain V which are infinitely refined. Correspondence (2.20) determines a mapping under which this sequence goes into a certain sequence of partitions $\{\Omega_i\}$ of the domain Ω, and since the maximal of the diameters of the subdomains V_i tends to zero, the same is with the maximal of the diameters of the subdomains Ω_i. The sequence of partitions of the domain Ω generates the corresponding sequence of integral sums each of which can be put down both in form (2.24) and (2.25). The limit of the sequence of integral sums of form (2.24) is equal to integral (2.23) while the limit of sums (2.25) is equal to integral (2.26). Thus, integrals (2.23) and (2.26) are the limits of the same integral sums and hence they coincide, that is

$$\iiint_{V} f(x, y, z)\,dv =$$

$$= \iiint_{\Omega} f(x(\xi, \eta, \zeta), \quad y(\xi, \eta, \zeta), \quad z(\xi, \eta, \zeta))\,|J(\xi, \eta, \zeta)|\,d\omega \tag{2.27}$$

Consequently, if there is a one-to-one continuous and continuously differentiable mapping of a bounded closed domain V onto a domain Ω, with a nonzero Jacobian, and if $f(x, y, z)$ is a continuous function defined in the domain V, formula (2.27) for changing variables in the triple integral is true.

We can easily show that the formula not only holds for a continuous function f but also for any bounded function continuous everywhere in V possibly except a set of points of volume zero.

Let us come back to formulas (2.20) determining a correspondence between the range V of the variables x, y, z and the range Ω of the variables ξ, η, ζ. The correspondence transforms an infinitesimal rectilinear parallelepiped

$$\xi_0 \leqslant \xi \leqslant \xi_0 + d\xi, \quad \eta_0 \leqslant \eta \leqslant \eta_0 + d\eta, \quad \zeta_0 \leqslant \zeta \leqslant \zeta_0 + d\zeta$$

of volume $d\omega = d\xi\,d\eta\,d\zeta$ lying in Ω into a curvilinear parallelepiped specified by the same inequalities and belonging to V. The

volume of the latter parallelepiped is equal to

$$dv = |J(\xi, \eta, \zeta)|\, d\xi\, d\eta\, d\zeta \qquad (2.28)$$

Thus, the absolute value $|J(\xi, \eta, \zeta)|$ of the Jacobian is the ratio of infinitesimal volumes corresponding to each other under mapping (2.20) (see Fig. 2.7).

In some simpler cases the Jacobian associated with a change of variables can be found on the basis of expression (2.28) for an element of volume, i.e. by means of purely geometric considerations

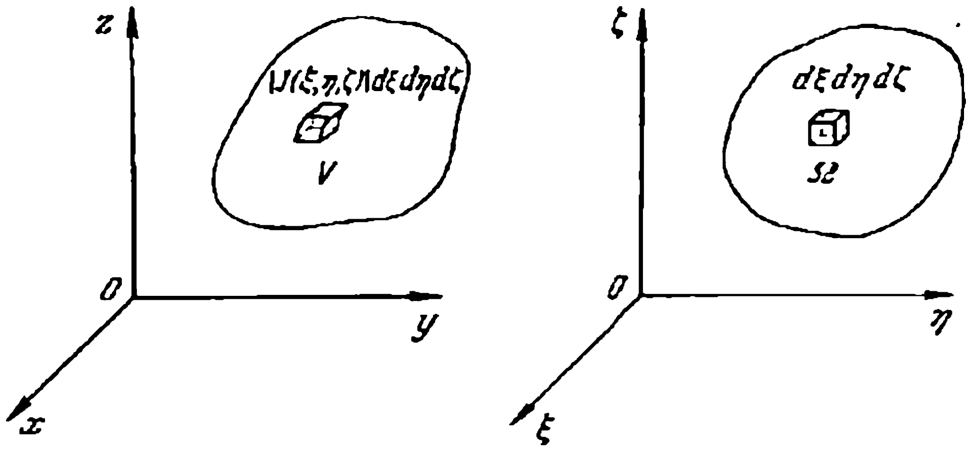


Fig. 2.7

without computing the corresponding derivatives and the determinant. We shall illustrate this technique by the examples of cylindrical and spherical coordinates.

Cylindrical coordinates. Consider an element of volume contained between three pairs of coordinate surfaces drawn infinitely close to each other, namely the cylinders of radii r and $r + dr$, two horizontal planes corresponding to certain values z and $z + dz$ of the

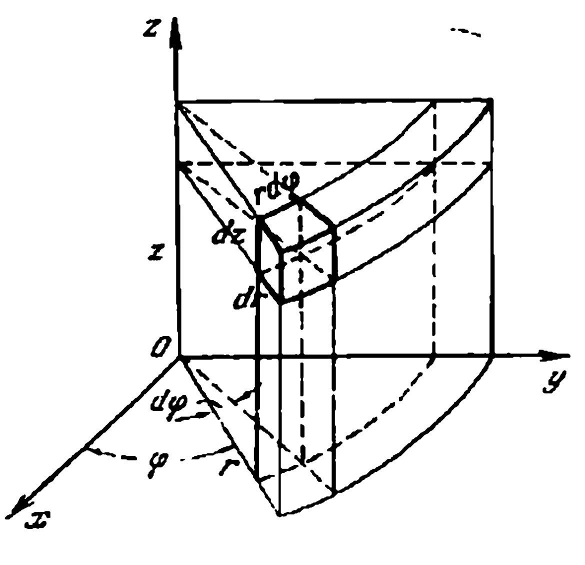


Fig. 2.8

coordinate z and two half-planes passing through the z-axis and forming angles φ and $\varphi + d\varphi$ with the x-axis. The volume element bounded by the surfaces is, to within infinitesimals of higher order

of smallness, a rectangular parallelepiped with edges dr, dz and $r\,d\varphi$ (see Fig. 2.8). Its volume is equal to

$$r\,dr\,d\varphi\,dz$$

which implies that the Jacobian of transformation of Cartesian coordinates into cylindrical ones is equal to r.

Spherical coordinates. Take a domain bounded by two spheres of radii r and $r + dr$, two semicones determined by certain angles θ and $\theta + d\theta$ (reckoned from the z-axis) and two half-planes forming

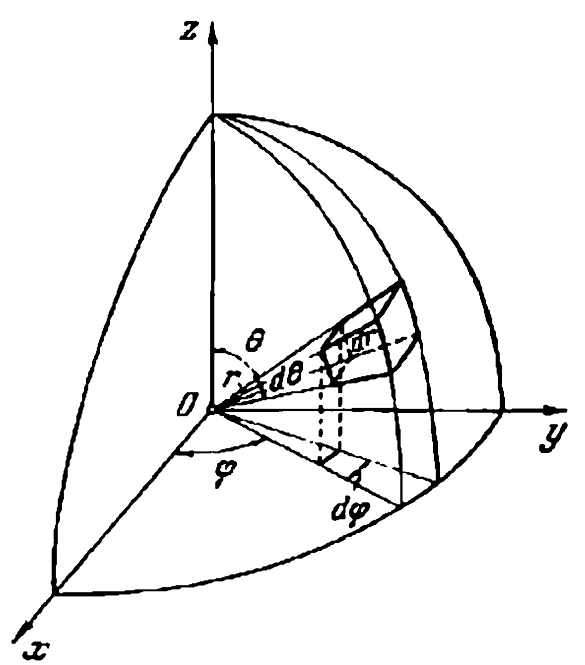


Fig. 2.9

angles φ and $\varphi + d\varphi$ with the x, z-plane. The domain can be thought of as a rectangular parallelepiped (to within infinitesimals of higher order) with edges $r\,d\theta$, dr and $r \sin\theta\,d\varphi$ (see Fig. 2.9). Consequently, its volume is equal to

$$r^2 \sin\theta\,dr\,d\theta\,d\varphi$$

which shows that the corresponding Jacobian is equal to

$$r^2 \sin\theta$$

§ 5. MULTIPLE INTEGRALS OF HIGHER ORDER

1. General Remarks. The definitions and facts which have been discussed in Chapter 1 for two variables, and in the present chapter for three variables, can be transferred to the case of an arbitrary number of variables. For this purpose we first of all define the volume of an n-dimensional parallelepiped.

As is well known from analytic geometry, the area of a parallelogram or the volume of a parallelepiped constructed on given vectors

is expressed by the formulas representing it as the absolute value of the determinant whose elements are the projections of the vectors on the coordinate axes. Starting from these results, we define the volume of an n-dimensional parallelepiped as the absolute value of the determinant with rows (or columns) formed of the coordinates of the vectors which are the edges of the parallelogram. Further, based on the volume of a parallelepiped, we can easily introduce the concept of volume for polyhedral n-dimensional figures and then define the volume for a wider class of domains lying in an n-dimensional space. After that the notion of an integral of a function $f(x_1, x_2, \ldots, x_n)$ dependent on arguments $x_1, x_2, \ldots, x_n$ is introduced as the limit of the corresponding integral sums. An n-fold multiple integral taken over an n-dimensional domain G is designated by the symbol

$$\iint\limits_G \ldots \int f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n$$

If the corresponding restrictions imposed on the domain G and on the integrand are fulfilled, the n-fold multiple integral can be reduced to an iterated integral in which n successive integrations with respect to each argument are performed:

$$\iint\limits_G \ldots \int f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n =$$

$$= \int\limits_a^b dx_1 \int\limits_{x_2^{(1)}(x_1)}^{x_2^{(2)}(x_1)} dx_2 \ldots \int\limits_{x_n^{(1)}(x_1, \ldots, x_{n-1})}^{x_n^{(2)}(x_1, \ldots, x_{n-1})} f(x_1, x_2, \ldots, x_n)\, dx_n$$

The formula for changing variables in the n-fold multiple integral is analogous to the corresponding formulas for double and triple integrals. Namely, if $x_i = x_i(y_1, y_2, \ldots, y_n)$ $(i = 1, 2, \ldots, n)$ the integral is transformed according to the formula

$$\int\limits_G \ldots \int f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n =$$

$$= \int\limits_\Gamma \ldots \int f(x_1(y_1, \ldots, y_n), \ldots, x_n(y_1, \ldots, y_n)) \times$$

$$\times \left| \frac{D(x_1, \ldots, x_n)}{D(y_1, \ldots, y_n)} \right| dy_1 \ldots dy_n$$

where Γ is the range of the variables $y_1, \ldots, y_n$.

2. Examples. All the basic facts of the theory of the double and triple integrals remain true for the n-dimensional case. Here we do not dwell in more detail on the general theory of n-fold multiple integrals and proceed to discuss some characteristic examples.

(1) *Gravitational attraction of two material bodies.* Although the real physical space we live in has only three dimensions there are various concrete problems involving multiple integrals of order higher than the third. As a simple example, let us derive the formula for the force of mutual gravitational attraction of two finite material solids. Let them occupy, respectively, domains G and G' and have volume densities $\rho(x, y, z)$ and $\rho'(x', y', z')$ (these material bodies lie in the same x, y, z-space but it is convenient to denote the coordinates of their points by different symbols). According to Newton's* law of universal gravitation, the projection on the x-axis of the force of attraction between two infinitesimal elements

$$dv = dx\,dy\,dz \quad \text{and} \quad dv' = dx'\,dy'\,dz'$$

of volume of the bodies is a quantity dF_x equal to

$$\gamma\frac{\rho(x, y, z)\,\rho'(x', y', z')}{r^3}(x-x')\,dx\,dy\,dz\,dx'\,dy'\,dz' \tag{2.29}$$

where γ is the constant of gravitation and

$$r=\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}$$

To find the total value of the projection F_x of the force of interaction between the bodies we must sum up expressions (2.29) over all the volume elements of both bodies. In other words, the projection F_x of the force of gravitational attraction between the bodies occupying the domains G and G' is equal to

$$\gamma\iiint\!\!\!\iiint_{G\times G'}\frac{\rho(x, y, z)\,\rho'(x', y', z')}{r^3}(x-x')\,dx\,dy\,dz\,dx'\,dy'\,dz' \tag{2.30}$$

The other two projections are expressed similarly. Here the point (x, y, z) runs throughout the domain G and the point (x', y', z') independently runs over the entire domain G'. Hence, integral (2.30) is taken over a domain in the six-dimensional space of the variables x, y, z, x', y', z'. Such a domain is usually denoted by the symbol $G \times G'$ and called the **(Cartesian) product of the domains** G **and** G'.

(2) Consider the integral

$$I_n=\iint\dots\int_G dx_1\,dx_2\dots dx_n \tag{2.31}$$

taken over the domain G determined by the inequalities

$$x_1\geqslant 0,\quad x_2\geqslant 0,\ \dots,\quad x_n\geqslant 0$$
$$x_1+x_2+\dots+x_n\leqslant 1$$

* Newton, Isaak (1642-1727), the great English mathematician and physicist.

Reducing integral (2.31) to an iterated integral we obtain

$$I_n = \iint_G \ldots \int dx_1\, dx_2 \ldots dx_n =$$

$$= \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \ldots \int_0^{1-x_1-x_2-\ldots-x_{n-1}} dx_n$$

Performing integration with respect to x_n and substituting th corresponding limits of integration we derive

$$I_n = \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \ldots \int_0^{1-x_1-\ldots-x_{n-2}} (1-x_1-\ldots-x_{n-1})\, dx_{n-1}$$

Next, integrating with respect to x_{n-1} and substituting the limit of integration we arrive at the formula

$$I_n = \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \ldots \int_0^{1-x_1-\ldots-x_{n-3}} \frac{(1-x_1-\ldots-x_{n-2})^2}{2!}\, dx_{n-2}$$

Continuing these successive integrations we finally obtain

$$I_n = \int_0^1 \frac{(1-x_1)^{n-1}}{(n-1)!}\, dx_1 = \frac{1}{n!}$$

(3) *The volume of an n-dimensional sphere.* An n-dimensional sphere (ball) of radius a with centre at the origin of coordinates is, by definition, the totality of all the points of the n-dimensional space whose coordinates satisfy the condition

$$x_1^2 + x_2^2 + \ldots + x_n^2 \leqslant a^2$$

The volume V_n of such a sphere is equal to the integral

$$\iint\limits_{x_1^2+x_2^2+\ldots+x_n^2 \leqslant a^2} \ldots \int dx_1\, dx_2 \ldots dx_n$$

The integral can be transformed in the following way. Putting $x_i = ay_i$ $(i = 1, 2, \ldots, n)$ we can write

$$V_n = a^n U_n$$

where U_n is the volume of unit n-dimensional sphere (of radius 1). Furthermore, since

$$U_n = \iint\limits_{x_1^2+x_2^2+\ldots+x_n^2\leqslant 1}\ldots\int dx_1\,dx_2\,\ldots\,dx_n =$$

$$= \int_{-1}^{1} dx_n \int\limits_{x_1^2+x_2^2+\ldots+x_{n-1}^2\leqslant 1-x_n^2}\ldots\int dx_1\ldots dx_{n-1} =$$

$$= \int_{-1}^{1} (1-x_n^2)^{\frac{n-1}{2}}\,dx_n \int\limits_{x_1^2+\ldots+x_{n-1}^2\leqslant 1}\ldots\int dx_1\ldots dx_{n-1} =$$

$$= U_{n-1}\int_{-1}^{1}(1-x_n^2)^{\frac{n-1}{2}}\,dx_n\,\text{*}$$

we can put $x_n = \cos\theta$ and thus receive

$$U_n = 2U_{n-1}\int_0^{\frac{\pi}{2}}\sin^n\theta\,d\theta \tag{2.32}$$

Now taking into account that $U_1 = 2$ (because one-dimensional unit sphere with centre at the origin is nothing but the line segment $[-1, 1]$, and the corresponding one-dimensional volume is its length) we can find, in succession, U_2, U_3 and so on.**

---

* Here, when computing the integral $\int\limits_{x_1^2+\ldots+x_{n-1}^2<1-x_n^2}\ldots\int dx_1\ldots dx_{n-1}$, we put $x_i=(1-x_n^2)^{\frac{1}{2}}y_i$ ($i = 1, 2, \ldots, n-1$) and thus obtain the expression

$$(1-x_n^2)^{\frac{n-1}{2}}\int\limits_{y_1^2+\ldots+y_{n-1}^2\leqslant 1}\ldots\int dy_1\ldots dy_{n-1} = (1-x_n^2)^{\frac{n-1}{2}}\int\limits_{x_1^2+\ldots+x_{n-1}^2\leqslant 1}\ldots\int dx_1\ldots dx_{n-1}.$$ — *Tr.*

** An explicit expression for U_n can be obtained with the help of so-called Euler's integrals (see Chapter 10, § 3 and, in particular, Example 3).

3 Elements of Differential Geometry

In this chapter we shall apply the differential and integral calculus to studying geometric objects, namely *curves* and *surfaces*. The division of mathematics in which various types of geometric configuration are investigated by means of mathematical analysis is called **differential geometry**.

In the framework of our course we can only present the fundamentals of differential geometry which is an extensive branch of mathematics closely related to mechanics, the theory of differential equations and other branches of knowledge.

§ 1. VECTOR FUNCTION OF A SCALAR ARGUMENT

1. Definition of a Vector Function. Limit. Continuity. It is convenient to define curves and surfaces by means of functions taking vectorial values (briefly, referred to as *vector functions*). Therefore we begin this chapter with a brief review of basic applications of mathematical analysis to vector functions. We shall not go into particulars because there are only a few facts here distinct from those of the theory of scalar functions.

Definition. Let, to each value of a variable t belonging to an interval $[a, b]$, there correspond a vector

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k} \tag{3.1}$$

Such vector is said to be a ***vector (vector-valued) function*** *of the scalar argument t.*

A vector function $\mathbf{r}(t)$ can be given the following visual interpretation. If the vectors $\mathbf{r}(t)$ corresponding to all the possible values of the argument t are laid off from the origin of coordinates their tips trace a curve (*the graph of the vector function*) which is called the **hodograph** of the function $\mathbf{r}(t)$ (Fig. 3.1). If the argument t is considered to be time the hodograph of the function $\mathbf{r}(t)$ is interpreted as the trajectory of motion of a point.

A constant vector

$$\mathbf{R} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

is said to be the limit of $\mathbf{r}(t)$ for $t \to t_0$ if

$$\lim_{t\to t_0} |\mathbf{r}(t) - \mathbf{R}| = 0 \tag{3.2}$$

where $|\mathbf{r}(t) - \mathbf{R}|$ is the length (absolute value, or modulus) of the vector $\mathbf{r}(t) - \mathbf{R}$. Condition (3.2) is equivalent to the three scalar relations

$$\lim_{t\to t_0} x(t) = a, \quad \lim_{t\to t_0} y(t) = b, \quad \lim_{t\to t_0} z(t) = c \tag{3.2'}$$

A vector function $\mathbf{r}(t)$ is called continuous at a point t_0 if

$$\lim_{t\to t_0} \mathbf{r}(t) = \mathbf{r}(t_0)$$

A vector function $\mathbf{r}(t)$ is continuous at a point t_0 if and only if its projections, i.e. the three scalar functions $x(t)$, $y(t)$ and $z(t)$, are continuous at t_0 (prove it). The sum, the difference, the scalar and vector products of continuous vector functions are again continuous (show it).

2. Differentiation of a Vector Function. A vector function $\mathbf{r}(t)$ is said to be differentiable at a point t if the limit

$$\lim_{\Delta t\to 0} \frac{\Delta \mathbf{r}(t)}{\Delta t} = \lim_{\Delta t\to 0} \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t)}{\Delta t}$$

exists. The limit is called the **derivative** of the vector function $\mathbf{r}(t)$ and is denoted by the symbols $\frac{d\mathbf{r}}{dt}$, $\mathbf{r}'(t)$ or $\dot{\mathbf{r}}(t)$. It can be easily proved that the existence of $\mathbf{r}'(t)$ is equivalent to the existence of the three derivatives $x'(t)$, $y'(t)$ and $z'(t)$, these quantities being connected by the relation

$$\mathbf{r}'(t) = x'(t)\,\mathbf{i} + y'(t)\,\mathbf{j} + z'(t)\,\mathbf{k}$$

The vector $\frac{\Delta \mathbf{r}}{\Delta t}$ is directed along the secant MM_1 of the hodograph of the function $\mathbf{r}(t)$ (see Fig. 3.2), and the vector $\frac{d\mathbf{r}}{dt}$ is in the direction of the limiting straight line to whose position the secant tends as the point M_1 approaches M. Hence, $\frac{d\mathbf{r}}{dt}$ goes along the tangent line to the hodograph at the point M.

From the point of view of kinematics, $\mathbf{r}'(t)$ is nothing but the velocity of a point whose law of motion is $\mathbf{r} = \mathbf{r}(t)$.

The following rules for differentiation of a vector function take place:

(1) if $\mathbf{r}(t) = \text{const}$ we have $\mathbf{r}'(t) \equiv 0$;

(2) $(k\mathbf{r}(t))' = k\mathbf{r}'(t)$ where $k = \text{const}$;

(3) $(u(t)\mathbf{r}(t))' = u'(t)\mathbf{r}(t) + u(t)\mathbf{r}'(t)$ where $u(t)$ is a scalar function;

(4) $(\mathbf{r}_1(t) \pm \mathbf{r}_2(t))' = \mathbf{r}_1'(t) \pm \mathbf{r}_2'(t)$;

(5) $(\mathbf{r}_1(t), \mathbf{r}_2(t))' = (\mathbf{r}_1'(t), \mathbf{r}_2(t)) + (\mathbf{r}_1(t), \mathbf{r}_2'(t))$;

(6) $[\mathbf{r}_1(t), \mathbf{r}_2(t)]' = [\mathbf{r}_1'(t), \mathbf{r}_2(t)] + [\mathbf{r}_1(t), \mathbf{r}_2'(t)]$ (it is necessary to preserve the order of the factors here);

(7) if $\mathbf{r} = \mathbf{r}(t)$ and $t = t(\tau)$, then

$$\frac{d\mathbf{r}}{d\tau} = \frac{d\mathbf{r}}{dt}\frac{dt}{d\tau}$$

which expresses the rule for differentiation of a composite vecto function.

The proof of the rules is left to the reader.

The following special cases of differentiation of a vector functio should be noted:

(a) *The derivative of a vector function of a constant direction.* Le a vector $\mathbf{r}(t)$ have an invariable direction in space (i.e. only it

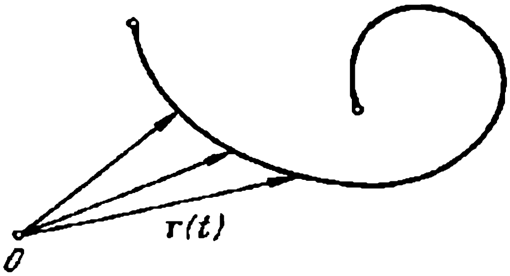


Fig. 3.1

length may depend on t). Then the vectors $\mathbf{r}(t)$ and $\mathbf{r}'(t)$ are colli near.

Actually, in this case the vector $\mathbf{r}(t)$ can be put down in the forn

$$\mathbf{r}(t) = u(t)\,\mathbf{e}$$

where $u(t)$ is a scalar function and $\mathbf{e}$ is a constant vector. The we have $\mathbf{r}'(t) = u'(t)\,\mathbf{e}$, that is $\mathbf{r}'(t) = \frac{u'(t)}{u(t)}\mathbf{r}(t)$.

(b) *The derivative of a vector function having a constant length.* If $|\mathbf{r}(t)| = \text{const}$ the vectors $\mathbf{r}(t)$ and $\mathbf{r}'(t)$ are orthogonal (perpendicular) to each other.

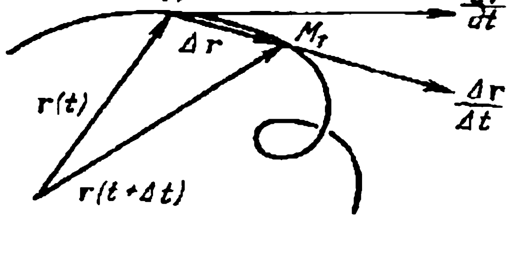


Fig. 3.2

Indeed, here we have $(\mathbf{r}(t), \mathbf{r}'(t)) = \text{const}$. After this relation has been differentiated we obtain

$$2(\mathbf{r}(t), \mathbf{r}'(t)) = 0, \quad \text{that is} \quad (\mathbf{r}(t), \mathbf{r}'(t)) = 0$$

which is what we set out to prove.

The geometric meaning of the last relation is quite clear. If $|\mathbf{r}(t)| = R$, the hodograph of the function $\mathbf{r}(t)$ entirely lies on the sphere of radius R with centre at the origin. The tangent to such a curve lies

in the tangent plane to the sphere and is therefore perpendicular to the radius vector $\mathbf{r}(t)$ joining the origin with the point of tangency.

The **differential of a vector function** $\mathbf{r}(t)$ is the vector

$$d\mathbf{r} = dx \cdot \mathbf{i} + dy \cdot \mathbf{j} + dz \cdot \mathbf{k}$$

In other words, we have

$$d\mathbf{r} = x'(t)\,dt \cdot \mathbf{i} + y'(t)\,dt \cdot \mathbf{j} + z'(t)\,dt \cdot \mathbf{k} = \mathbf{r}'(t)\,dt$$

which means that the differential of a vector function is equal to the product of its derivative by the differential (i.e. increment) of the independent variable. As in the case of a scalar function, the differential $d\mathbf{r}$ of a vector function differs from its increment $\Delta\mathbf{r}$ by a quantity of an order of smallness higher than the first relative to Δt.

3. Hodograph. Singular Points. We have defined the hodograph of a vector function $\mathbf{r}(t)$ as a curve which is traced by the terminal of the vector $\mathbf{r}(t)$, when t varies, if its tail is always kept at a fixed point.

As has been shown, if $\mathbf{r}(t)$ is a differentiable vector function the vector $\mathbf{r}'(t)$ is directed along the tangent to the hodograph at all points where $\mathbf{r}'(t) \neq 0$. The points at which the derivative $\mathbf{r}'(t)$

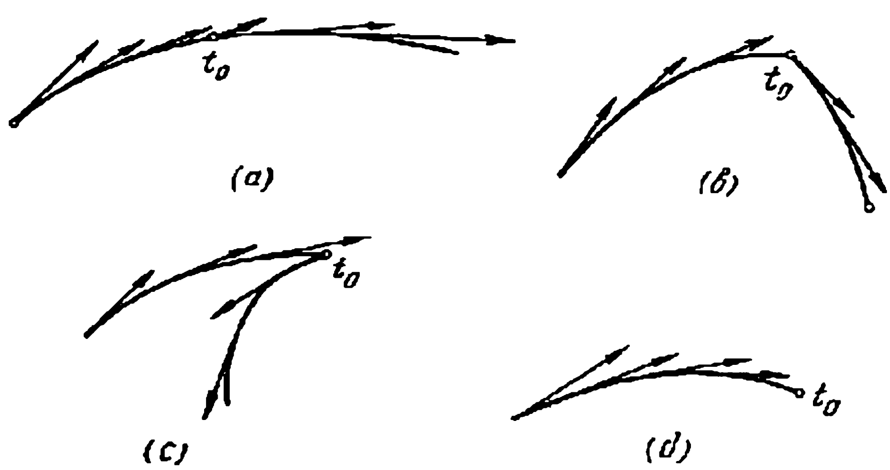


Fig. 3.3

does not exist or vanishes are referred to as **singular points**. Here we give several examples of singular points (see Fig. 3.3). In the motion of a point according to a law $\mathbf{r} = \mathbf{r}(t)$ the trajectory may be a "smooth" curve but the velocity $\mathbf{v}(t) = \mathbf{r}'(t)$ may tend to zero as $t \to t_0$. Then the material point is in an instantaneous state of rest at the point $\mathbf{r}(t_0)$ at the moment $t = t_0$. A singularity of this kind characterizes the motion itself but not the geometric curve along which the point moves (Fig. 3.3*a*). In some other cases this can be followed by a change in the direction of motion (i.e. the trajectory may be broken at the point $\mathbf{r}(t_0)$; see Fig. 3.3*b*). In this

case we have a singularity both of the motion and of the corresponding geometric curve. It may turn out that the trajectory is broken at a point $\mathbf{r}(t_0)$ but the velocity $\mathbf{r}'(t)$ does not tend to zero in approaching the point (see again Fig. 3.3*b*). In this case the moving point $\mathbf{r}(t)$ is subjected to an impact (when it passes through the point $\mathbf{r}(t_0)$) which instantaneously changes its velocity in a jump-like fashion. Further, the trajectory of motion can have a cusp (Fig. 3.3*c*) and then the velocity of the material point may either tend to zero in the vicinity of the cusp or change jump-like. Finally, the function $\mathbf{r}(t)$ may turn into zero for $t = t_0$ without assuming nonzero values for the subsequent values of $t > t_0$. Then the point has a state of rest at the end of the motion (Fig. 3.3*d*). These and many other singularities of motion may be of interest for studying some concrete cases but at the same time they can rarely be treated by means of the general methods. Therefore in what follows we shall exclude such singularities from our discussion and consider the motions for which $\mathbf{r}'(t)$ exists everywhere and does not vanish.

4. Taylor's Formula. For a vector function we have Taylor's formula

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{r}'(t)\,\Delta t + \frac{1}{2}\,\mathbf{r}''(t)\,\Delta t^2 + \ldots + \frac{1}{n!}\left(\mathbf{r}^{(n)}(t) + \boldsymbol{\alpha}\right)\Delta t^n \tag{3.3}$$

where $\boldsymbol{\alpha}$ is a vector which tends to zero as $\Delta t \to 0$. In fact, applying Taylor's formula to each projection $x(t)$, $y(t)$ and $z(t)$* of the vector $\mathbf{r}(t)$ we obtain the relation

$$x(t+\Delta t) = x(t) + x'(t)\,\Delta t + \frac{1}{2}\,x''(t)\,\Delta t^2 + \ldots + \frac{1}{n!}\left(x^{(n)}(t) + \alpha_1\right)\Delta t^n$$

for $x(t)$ and two similar relations for $y(t)$ and $z(t)$. Multiplying these relations, respectively, by $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ and adding up the results we arrive at formula (3.3).

Thus we see that most of the basic notions and rules of differential calculus are transferred without essential changes from scalar functions to the vector-valued ones. But it should be noted that such conclusions cannot be drawn automatically because there are exceptions to the rule. For instance, the well known theorem on finite increments (Lagrange's theorem; e.g. see [8], Chapter 8, § 9) is not true for vector functions. (Let the reader construct an illustrative example.)

5. Integral of a Vector Function with Respect to Scalar Argument. As in the case of a scalar function, we can form integral sums for

---

* Here we suppose, of course, that each projection $x(t)$, $y(t)$ and $z(t)$ of the vector function $\mathbf{r}(t)$ satisfies the conditions for the applicability of Taylor's formula to it.

a vector function $\mathbf{r}(t)$ defined on an interval $a \leqslant t \leqslant b$ and consider their limit as the maximal length of the segments forming the partition of the interval $[a, b]$ tends to zero. The limit is referred to as the **integral of $\mathbf{r}(t)$ over the interval $[a, b]$** and is designated as

$$\int_a^b \mathbf{r}(t)\,dt$$

By analogy with scalar functions, we can easily prove that the limit (i.e. the integral) is sure to exist if $\mathbf{r}(t)$ is continuous on $[a, b]$.

The existence of the limit of a vector integral sum

$$\sum_{i=1}^{n} \mathbf{r}(\tau_i)(t_i - t_{i-1})$$

(where $a = t_0 < t_1 < \ldots < t_n = b$, $t_{i-1} \leqslant \tau_i \leqslant t_i$) is obviously equivalent to the existence of the limits of the three scalar integral sums corresponding to the projections $x(t)$, $y(t)$ and $z(t)$ of the vector function $\mathbf{r}(t)$, and we have

$$\int_a^b \mathbf{r}(t)\,dt = \mathbf{i}\cdot\int_a^b x(t)\,dt + \mathbf{j}\cdot\int_a^b y(t)\,dt + \mathbf{k}\cdot\int_a^b z(t)\,dt$$

The ordinary properties of the integrals of scalar functions are easily extended to the case of the integrals of vector functions. For example, we have

$$\int_a^b u'(t)\,\mathbf{r}(t)\,dt = u(b)\,\mathbf{r}(b) - u(a)\,\mathbf{r}(a) - \int_a^b u(t)\,\mathbf{r}'(t)\,dt$$

which is the formula of integration by parts where $u(t)$ is a scalar function. The formulas connecting integration with basic operations of vector algebra are also easily established. For instance,

$$\int_a^b [\mathbf{c}, \mathbf{r}(t)]\,dt = \left[\mathbf{c}, \int_a^b \mathbf{r}(t)\,dt\right]$$

where $\mathbf{c}$ is a constant vector.

6. Vector Functions of Several Scalar Arguments. We can also consider vector functions dependent not on one but on many scalar arguments (in particular, we shall encounter vector functions of two scalar arguments in the present chapter in the study of surfaces). The concept of a partial derivative and other concepts of analysis are easily generalized to these functions.

§ 2. SPACE CURVES

1. Vector Equation of a Curve. Vector functions of a scalar argument provide a convenient method of determining space curves.* Indeed, suppose we are given a continuous vector function $\mathbf{r}(t)$ $(a \leqslant t \leqslant b)$. Then, after constructing its hodograph, we obtain a space curve γ. Conversely, if a space curve γ is defined in a certain way we can try to determine it by means of a vector function.

We say that a curve γ is *represented parametrically* if there is a *one-to-one* correspondence which attributes a certain value of a parameter t belonging to an interval $[a, b]$ to each point of the curve, the correspondence being *continuous* at each point of the interval.** The latter condition means that the distance between the points $\mathbf{r}(t_0)$ and $\mathbf{r}(t)$ of the curve tends to zero if $t \to t_0$. If a curve γ is represented parametrically the radius vector of each of its points is uniquely determined by the corresponding value of the parameter t, that is

$$\mathbf{r} = \mathbf{r}(t) \quad (\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) \tag{3.4}$$

Relation (3.4) is referred to as a **parametric (vector) equation** of the curve γ. Vector equation (3.4) can apparently be replaced by the three scalar parametric equations

$$x = x(t), \quad y = y(t), \quad z = z(t)$$

Applying the terminology of § 1 we can say that a parametric equation of a curve provides its representation as the hodograph of a vector function $\mathbf{r}(t)$.

In what follows we shall only consider the curves and their parametric representations for which the corresponding vector functions $\mathbf{r}(t)$ are triply continuously differentiable.

Example. Let us put

$$\mathbf{r}(t) = \mathbf{i}a \cos t + \mathbf{j}\, a \sin t + \mathbf{k}\, bt \tag{3.5}$$

This parametric equation determines a curve called a **screw line (circular helix**; see Fig. 3.4).

When considering a curve we can introduce its parametric representation in various ways. For instance, if a curve γ is given by an equation $\mathbf{r} = \mathbf{r}(t)$, $a \leqslant t \leqslant b$, we can put

$$t = t(\tau), \quad \alpha \leqslant \tau \leqslant \beta$$

where $t(\tau)$ is a monotone function such that $t'(\tau) > 0$, $t(\alpha) = a$ and $t(\beta) = b$, and regard τ as a new parameter providing the equation $\mathbf{r} = \mathbf{r}(t(\tau))$ for the curve γ.

---

* We do not specify the notion of a curve here. A discussion concerning this question can be found, for instance, in [8], Chapter 11, § 1.

** The condition that the correspondence is one-to-one means that we are considering the curves without points of self-intersection.

In many cases it is convenient to take as a parameter the arc length of the curve reckoned from a fixed point. The transition from an arbitrary parameter entering into a parametric representation of a curve to the arc length of the curve can be performed as follows.

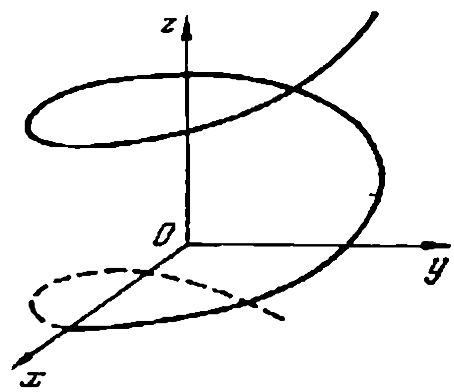


Fig. 3.4

Let γ be a curve and t a parameter entering into its parametric equation. Choose a point M_0 on γ which corresponds to a certain value $t = t_0$ of the parameter and consider it to be an initial point. Next take an arbitrary point M on γ. The length l of the arc M_0M is expressed by the well known formula

$$l = \int_{t_0}^{t} \sqrt{x'^2 + y'^2 + z'^2}\, dt^*, \quad \text{i.e.} \quad l = \int_{t_0}^{t} |\mathbf{r}'(t)|\, dt$$

where t is the value of the parameter corresponding to the point M. The formula determines l as a single-valued and continuous function of t: $l = l(t)$. If the function $\mathbf{r}(t)$ is such that $\mathbf{r}'(t)$ does not vanish anywhere we have $l'(t) \neq 0$ at all the points and consequently t can be represented as a single-valued continuous function of l: $t = t(l)$. (On the existence of an inverse function of type $t = t(l)$ see, for instance, [8], Chapter 11, § 1.) Now putting $\mathbf{r} = \mathbf{r}(t(l))$ we thus represent $\mathbf{r}$ as a function of arc length l, i.e. obtain a parametric equation of the curve in which the arc length serves as a parameter.

Example. Consider again circular helix (3.5). We have

$$dl = \sqrt{a^2 \sin^2 t + a^2 \cos^2 t + b^2}\, dt = \sqrt{a^2 + b^2}\, dt$$

---

* The formula means in fact that the curve $x = x(t)$, $y = y(t)$, $z = z(t)$ is regarded as being a "broken line" with an infinite number of infinitesimal segments (dx, dy, dz). The length of a single segment is given by the Pythagorean theorem and is equal to

$$\sqrt{(dx)^2 + (dy)^2 + (dz)^2} = \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}\, dt$$

The "sum" of the lengths of the "segments", that is the integral

$$\int_{t_0}^{t} \sqrt{x'^2 + y'^2 + z'^2}\, dt$$

is just equal to the length of the curve.

for it and hence $l=\sqrt{a^2+b^2}\,t$. Passing to the parameter l we can rewrite the equation of the circular helix in the form

$$\mathbf{r}(l)=\mathbf{i}\,a\cos\frac{l}{\sqrt{a^2+b^2}}+\mathbf{j}\,a\sin\frac{l}{\sqrt{a^2+b^2}}+\mathbf{k}\,b\frac{l}{\sqrt{a^2+b^2}}$$

Note. If a parameter t entering into an equation

$$\mathbf{r}=\mathbf{r}(t)$$

is thought of as time the curve determined by the equation can be regarded as the trajectory of a point moving from an initial position with the velocity $\mathbf{r}'(t)$. But a point can be in various motions along the same curve because when specifying a curve we only prescribe the direction of the velocity at each moment but not its numerical value. In particular, we can consider the case when the velocity $\mathbf{r}'$ of motion is all the time identically equal to unity in its modulus. It is just the case when the arc length l is taken as the parameter t. Indeed, we have $d\mathbf{r}=\mathbf{i}\,dx+\mathbf{j}\,dy+\mathbf{k}\,dz$ and consequently

$$\left|\frac{d\mathbf{r}}{dl}\right|=\frac{\sqrt{(dx)^2+(dy)^2+(dz)^2}}{dl}=\frac{dl}{dl}=1 \tag{3.6}$$

Thus, from the point of view of kinematics, various possible ways of parametric representation of a curve can be interpreted as the corresponding laws of motion of particles tracing the same trajectory with different velocities. Then an equation of the form

$$\mathbf{r}=\mathbf{r}(l)$$

where l is arc length describes the motion of a particle with unit (in its modulus) velocity.

2. Moving Trihedron. Consider a curve given by an equation

$$\mathbf{r}=\mathbf{r}(l) \tag{3.7}$$

At each point M of the curve (corresponding to a value l), the unit vector

$$\boldsymbol{\tau}=\dot{\mathbf{r}}(l)^*$$

determines the direction of the tangent to the curve. The vector

$$\ddot{\mathbf{r}}=\dot{\boldsymbol{\tau}}$$

is orthogonal to $\boldsymbol{\tau}$ (because it is the derivative of the vector $\boldsymbol{\tau}$ having constant length; see Sec. 2 in § 1). After the vector $\ddot{\mathbf{r}}$ has been divided

* Here and henceforward we denote the derivatives of $\mathbf{r}$ with respect to arc length by the symbols $\dot{\mathbf{r}}$, $\ddot{\mathbf{r}}$ etc. and use the notation $\mathbf{r}'$, $\mathbf{r}''$ etc. for the derivatives with respect to an arbitrary parameter.

by $|\ddot{\mathbf{r}}|$ we arrive at the unit vector

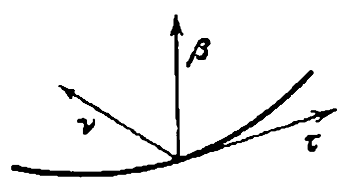
$$\mathbf{v}=\frac{\ddot{\mathbf{r}}}{|\ddot{\mathbf{r}}|} \tag{3.8}$$

orthogonal to $\boldsymbol{\tau}$.* Furthermore, let us take the vector

$$\boldsymbol{\beta}=[\boldsymbol{\tau}, \mathbf{v}] \tag{3.9}$$

where the square brackets designate the vector product. The vectors $\boldsymbol{\tau}$, $\mathbf{v}$ and $\boldsymbol{\beta}$ form a triad of mutually orthogonal unit vectors which is referred to as the **moving (natural) trihedron** of curve (3.7) at

Fig. 3.5



the point M (Fig. 3.5). The trihedron is thought of as being rigidly connected with the curve at each of its points and therefore the shape of the curve can be completely characterized by describing the motion of the trihedron in space as its vertex moves along the curve.

It should be noted that besides relation (3.9) the vectors $\boldsymbol{\tau}$, $\mathbf{v}$, $\boldsymbol{\beta}$ satisfy the two similar relations

$$[\mathbf{v}, \boldsymbol{\beta}]=\boldsymbol{\tau}, \quad [\boldsymbol{\beta}, \boldsymbol{\tau}]=\mathbf{v}$$

The vectors $\boldsymbol{\tau}$, $\mathbf{v}$ and $\boldsymbol{\beta}$ determine, respectively, the directions of the straight lines called the **tangent**, the **normal (principal normal)** and the **binormal** to the curve at the corresponding point. The vectors are referred to as *unit vectors in the direction of the tangent, principal normal and binormal.*

3. Frenet-Serret Formulas. The motion of the moving trihedron of a curve is specified by the velocities characterizing the rate of change of the vectors $\boldsymbol{\tau}$, $\mathbf{v}$ and $\boldsymbol{\beta}$, that is by the derivatives of the vectors with respect to l. Let us find the derivatives.

We have already dealt with the derivative of the vector $\boldsymbol{\tau}$ which is the vector $\ddot{\mathbf{r}}$. Introducing the notation

$$k=|\ddot{\mathbf{r}}|$$

* The vector $\mathbf{v}$ is not defined for the points where $\ddot{\mathbf{r}}=0$. Such points (called points of rectification) will be excluded from our consideration.

we rewrite the derivative in the form

$$\dot{\boldsymbol{\tau}} = k\mathbf{v}$$

where k is a nonnegative number.

We now consider the vector $\boldsymbol{\beta}$. Its derivative $\dot{\boldsymbol{\beta}}$ is perpendicular to it as a derivative of unit vector. Furthermore, it is also perpendicular to $\boldsymbol{\tau}$. In fact, we have $\boldsymbol{\beta} = [\boldsymbol{\tau}, \mathbf{v}]$ and hence $\dot{\boldsymbol{\beta}} = [\dot{\boldsymbol{\tau}}, \mathbf{v}] + [\boldsymbol{\tau}, \dot{\mathbf{v}}] = [k\mathbf{v}, \mathbf{v}] + [\boldsymbol{\tau}, \dot{\mathbf{v}}] = [\boldsymbol{\tau}, \dot{\mathbf{v}}]$, the last vector being perpendicular to $\boldsymbol{\tau}$. The vector $\dot{\boldsymbol{\beta}}$ is perpendicular to $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$ and thus collinear to $\mathbf{v}$. Consequently, we can put

$$\dot{\boldsymbol{\beta}} = -\varkappa\mathbf{v}$$

where $\varkappa$ is a numerical coefficient.*

Finally, compute $\dot{\mathbf{v}}$. We have

$$\dot{\mathbf{v}} = \frac{d}{dl}[\boldsymbol{\beta}, \boldsymbol{\tau}] = [\dot{\boldsymbol{\beta}}, \boldsymbol{\tau}] + [\boldsymbol{\beta}, \dot{\boldsymbol{\tau}}] = [-\varkappa\mathbf{v}, \boldsymbol{\tau}] + [\boldsymbol{\beta}, k\mathbf{v}] = -k\boldsymbol{\tau} + \varkappa\boldsymbol{\beta}$$

Thus, we have obtained the following formulas for the derivatives $\dot{\boldsymbol{\tau}}$, $\dot{\mathbf{v}}$ and $\dot{\boldsymbol{\beta}}$:

$$\dot{\boldsymbol{\tau}} = k\mathbf{v} \qquad (3.10)$$

$$\dot{\mathbf{v}} = -k\boldsymbol{\tau} + \varkappa\boldsymbol{\beta} \qquad (3.11)$$

$$\dot{\boldsymbol{\beta}} = -\varkappa\mathbf{v} \qquad (3.12)$$

They are known as the **Frenet-Serret**** **formulas**. The formulas involve two scalar quantities, namely k and $\varkappa$. The quantity k is called the **curvature** (the **first curvature**) of the curve and $\varkappa$ is called its **torsion** (or the **second curvature**). The geometric meaning of the curvature and the torsion will be discussed later. The reciprocals of k and $\varkappa$ are referred to, respectively, as the **radii of curvature** and **torsion**.

4. Evaluating Curvature and Torsion. We have, by definition,

$$k = |\ddot{\mathbf{r}}| \qquad (3.13)$$

Therefore, to compute the curvature of a curve $\mathbf{r} = \mathbf{r}(l)$ it is sufficient to find the vector $\ddot{\mathbf{r}}(l)$ and determine its length.

To find the torsion $\varkappa$ we take the equalities

$$\dot{\mathbf{r}} = \boldsymbol{\tau} \quad \text{and} \quad \ddot{\mathbf{r}} = k\mathbf{v}$$

---

* The coefficient $\varkappa$ can be positive, negative or zero. We use the notation $\varkappa$ instead of $\varkappa$ because this will be convenient for our further aims.

** Frenet, Jean Frederic (1816-1900) and Serret, Joseph Alfred (1819-1885), French mathematicians.

and differentiate the latter once again with respect to l. Applying formula (3.11) for $\dot{\mathbf{v}}$ we thus obtain

$$\dddot{\mathbf{r}} = k\dot{\mathbf{v}} - k^2\boldsymbol{\tau} + k\varkappa\boldsymbol{\beta}$$

The last three relations imply that

$$(\dot{\mathbf{r}}, \ddot{\mathbf{r}}, \dddot{\mathbf{r}}) = k^2\varkappa \tag{3.14}$$

whence $\varkappa = \frac{(\dot{\mathbf{r}}, \ddot{\mathbf{r}}, \dddot{\mathbf{r}})}{k^2}$, i.e.

$$\varkappa = \frac{(\dot{\mathbf{r}}, \ddot{\mathbf{r}}, \dddot{\mathbf{r}})}{|\ddot{\mathbf{r}}|^2} \tag{3.15}$$

Formulas (3.13) and (3.15) enable us to compute the curvature k and the torsion $\varkappa$ if the arc length has been chosen as the parameter. In case a curve is given by an equation

$$\mathbf{r} = \mathbf{r}\,(t)$$

where $\mathbf{r}\,(t)$ is a triply continuously differentiable function of an arbitrary parameter t we can regard t as being a function of arc length l which yields

$$\begin{gathered} \frac{d\mathbf{r}}{dl} = \frac{d\mathbf{r}}{dt}\cdot\frac{dt}{dl}, \quad \frac{d^2\mathbf{r}}{dl^2} = \frac{d^2\mathbf{r}}{dt^2}\left(\frac{dt}{dl}\right)^2 + \frac{d\mathbf{r}}{dt}\frac{d^2t}{dl^2}, \\ \frac{d^3\mathbf{r}}{dl^3} = \frac{d^3\mathbf{r}}{dt^3}\left(\frac{dt}{dl}\right)^3 + 3\,\frac{d^2\mathbf{r}}{dt^2}\,\frac{dt}{dl}\,\frac{d^2t}{dl^2} + \frac{d\mathbf{r}}{dt}\,\frac{d^3t}{dl^3} \end{gathered} \tag{3.16}$$

The first relation can be rewritten as

$$\boldsymbol{\tau} = \mathbf{r}'\,\frac{dt}{dl}$$

It follows that

$$\frac{dt}{dl} = \frac{1}{|\mathbf{r}'\,(t)|} \tag{3.17}$$

because $|\boldsymbol{\tau}| = 1$ $\left(\text{here we assume that } t \text{ and } l \text{ vary in the same direction, i.e. } \frac{dt}{dl} > 0\right)$. Further, taking the vector product of the first two equalities (3.16) we derive

$$\left[\frac{d\mathbf{r}}{dl}, \frac{d^2\mathbf{r}}{dl^2}\right] = \left[\frac{d\mathbf{r}}{dt}, \frac{d^2\mathbf{r}}{dt^2}\right]\left(\frac{dt}{dl}\right)^3$$

or, since $\left[\frac{d\mathbf{r}}{dl}, \frac{d^2\mathbf{r}}{dl^2}\right] = k\boldsymbol{\beta}$,

$$k\boldsymbol{\beta} = [\mathbf{r}'\,(t), \mathbf{r}''\,(t)]\left(\frac{dt}{dl}\right)^3 \tag{3.18}$$

Since $|\beta| = 1$, it follows from (3.17) and (3.18) that

$$k = \frac{|[\mathbf{r}'(t),\ \mathbf{r}''(t)]|}{|\mathbf{r}'(t)|^3} \tag{3.19}$$

Substituting expression (3.16) into relation (3.14) we receive

$$(\mathbf{r}'(t),\ \mathbf{r}''(t),\ \mathbf{r}'''(t)) \left(\frac{dt}{dl}\right)^6 = k^2 \varkappa \tag{3.20}$$

The last two equalities imply the final formula for the torsion:

$$\varkappa = \frac{(\mathbf{r}'(t),\ \mathbf{r}''(t),\ \mathbf{r}'''(t))}{|[\mathbf{r}'(t),\ \mathbf{r}''(t)]|^2} \tag{3.21}$$

Exercise. Find the curvature and the torsion of the screw line

$$\mathbf{r} = \mathbf{i}\, a \cos t + \mathbf{j}\, a \sin t + \mathbf{k}\, bt$$

Note. Let us come back to formulas (3.16). They indicate that the vectors $\mathbf{r}'$ and $\mathbf{r}''$ are linearly expressible in terms of the vectors $\dot{\mathbf{r}}$ and $\ddot{\mathbf{r}}$. In other words, the vectors $\mathbf{r}'$ and $\mathbf{r}''$ lie in the same plane as the vectors $\dot{\mathbf{r}}$ and $\ddot{\mathbf{r}}$. The plane is called the **osculating plane**. Hence, the osculating plane of a curve (at a given point) can be defined as a plane containing the vectors $\mathbf{r}'(t)$ and $\mathbf{r}''(t)$ (irrespective of the specific choice of the parameter). If t is interpreted as time and the equation

$$\mathbf{r} = \mathbf{r}(t)$$

as the law of motion of a point we can say that the osculating plane is the one that contains the velocity vector and the acceleration vector.

5. Coordinate System Connected with Moving Trihedron. For a curve $\mathbf{r}(l)$, the three vectors $\boldsymbol{\tau}$, $\boldsymbol{\nu}$ and $\boldsymbol{\beta}$ specify a coordinate system (for which they are the base vectors) at each point M of the curve, the system varying, in the general case, as the point moves along the curve. The axes of such a coordinate system are:

(1) the *tangent* (its direction is determined by the vector $\boldsymbol{\tau}$),

(2) the *principal normal* (its direction coincides with that of the vector $\boldsymbol{\nu}$),

(3) the *binormal* (which goes along the vector $\boldsymbol{\beta}$).

The coordinate planes of the system are:

(1) the plane drawn through the point M perpendicularly to $\boldsymbol{\tau}$ (i.e. the plane containing the principal normal and the binormal); it is called the **normal plane** to the curve $\mathbf{r} = \mathbf{r}(l)$ at the point M.

(2) the plane passing through the point M perpendicularly to $\boldsymbol{\nu}$ which is referred to as the **rectifying plane**.

(3) the plane passing through the point M and perpendicular to β (that is the plane in which $\dot{\mathbf{r}}$ and $\ddot{\mathbf{r}}$ lie). This is the osculating plane we have already dealt with.

The disposition of these straight lines (axes) and planes is depicted in Fig. 3.6.

Problem. Write the equations of the tangent, principal normal and binormal for a curve $\mathbf{r} = \mathbf{r}\,(l)$, and also the equations of the normal, rectifying and osculating planes, at a point $\mathbf{r}_0 = \mathbf{r}\,(l_0)$.

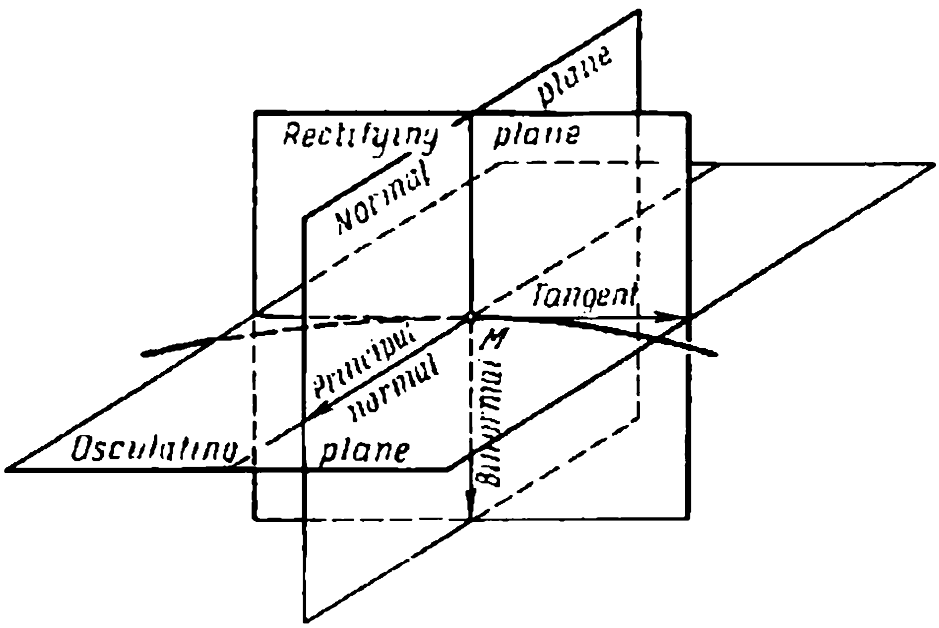


Fig. 3.6

Solution. The vector equation of a straight line passing through a point with radius vector $\mathbf{r}_0$ in the direction of a vector $\mathbf{a}$ is written as

$$\rho = \mathbf{r}_0 + \lambda \mathbf{a}, \quad -\infty < \lambda < \infty$$

where ρ is the radius vector of the moving point and λ is a parameter; the equation of a plane drawn through a point with radius vector $\mathbf{r}_0$ perpendicularly to a vector $\mathbf{n}$ is of the form

$$(\rho - \mathbf{r}_0, \mathbf{n}) = 0$$

This immediately implies the following equations:

$\rho = \mathbf{r}_0 + \lambda \dot{\mathbf{r}}_0$ (*tangent line*);

$\rho = \mathbf{r}_0 + \lambda \ddot{\mathbf{r}}_0$ (*principal normal*);

$\rho = \mathbf{r}_0 + \lambda\, [\dot{\mathbf{r}}_0, \ddot{\mathbf{r}}_0]$ (*binormal*);

$(\rho - \mathbf{r}_0, \dot{\mathbf{r}}_0) = 0$ (*normal plane*);

$(\rho - \mathbf{r}_0, \ddot{\mathbf{r}}_0) = 0$ (*rectifying plane*);

$(\rho - \mathbf{r}_0, [\dot{\mathbf{r}}_0, \ddot{\mathbf{r}}_0]) = 0$ (*osculating plane*)

where $\mathbf{r}_0 = \mathbf{r}\,(l_0)$, $\dot{\mathbf{r}}_0 = \dot{\mathbf{r}}\,(l_0)$ and $\ddot{\mathbf{r}}_0 = \ddot{\mathbf{r}}\,(l_0)$.

Exercises

1. Write the equations of the tangent, principal normal and binormal to the curve

$$\mathbf{r} = \mathbf{r}(t)$$

Hint. Note that the vector $[\mathbf{r}', \mathbf{r}'']$ is in the direction of the binormal and the vector $[\mathbf{r}', [\mathbf{r}', \mathbf{r}'']]$ goes along the principal normal.

2. Write the equations of the normal, rectifying and osculating planes for a curve $\mathbf{r} = \mathbf{r}(t)$.

3. Put down the equations of the tangent, principal normal and binormal and also of the normal, osculating and rectifying planes for the circular helix

$$x = a\cos t, \quad y = a\sin t, \quad z = bt$$

at the point $t = 0$.

6. The Shape of a Curve in the Vicinity of Its Point. To investigate the shape of a curve in the vicinity of its point we shall take advantage of the coordinate system determined by the moving trihedron of the curve.

Suppose the derivatives

$$\dot{\mathbf{r}}_0 = \dot{\mathbf{r}}(l_0), \quad \ddot{\mathbf{r}}_0 = \ddot{\mathbf{r}}(l_0) \text{ and } \dddot{\mathbf{r}}_0 = \dddot{\mathbf{r}}(l_0)$$

are different from zero at a point $\mathbf{r}_0 = \mathbf{r}(l_0)$ and expand the function $\mathbf{r}(l)$ in a neighbourhood of the point l_0 by means of Taylor's formula:

$$\mathbf{r}(l) = \mathbf{r}_0 + \dot{\mathbf{r}}_0\Delta l + \frac{1}{2}\ddot{\mathbf{r}}_0\Delta l^2 + \frac{1}{6}\dddot{\mathbf{r}}_0\Delta l^3 + O(\Delta l^4)^{*}, \quad \Delta l = l - l_0 \qquad (3.22)$$

Now take the coordinate system specified by the moving trihedron, i.e. choose the point $\mathbf{r}_0$ as the origin of coordinates and the tangent, principal normal and binormal lines as the x-axis, y-axis and z-axis, respectively. By applying the Frenet-Serret formulas for computing the derivatives $\ddot{\mathbf{r}}$ and $\dddot{\mathbf{r}}$ we can substitute the following three scalar equalities

$$x = \Delta l - \frac{k^2}{6}\Delta l^3 + O(\Delta l^4) \qquad (3.23a)$$

$$y = \frac{1}{2}k\Delta l^2 + \frac{\dot{k}}{6}\Delta l^3 + O(\Delta l^4) \qquad (3.23b)$$

$$z = \frac{1}{6}k\varkappa\Delta l^3 + O(\Delta l^4) \qquad (3.23c)$$

* The symbol $O(\Delta l^4)$ designates a quantity of the order of Δl^4.

for vector relation (3.22). Let us investigate the projections of the curve on the osculating and the rectifying planes.

We take equalities (3.23a) and (3.23b) and limit ourselves to the principal terms. The equalities then take the form

$$x=\Delta l, \quad y=\frac{1}{2}k\Delta l^2$$

Eliminating Δl from these relations we obtain the equation of a parabola (Fig. 3.7):

$$y=\frac{1}{2}kx^2$$

which is, to within the terms of the order of Δl^3, the projection of the curve $\mathbf{r}=\mathbf{r}(l)$ on the osculating plane. The curvature k being, by definition, positive, the parabola opens upwards or downwards

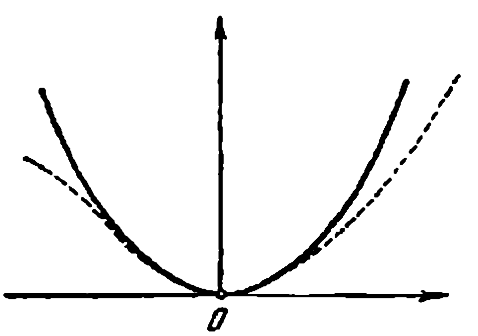


Fig. 3.7

according to the unit vector $\mathbf{v}$, and the greater k, the greater the rate at which the branches of the parabola are turning out of the tangent in the direction of the vector $\mathbf{v}$.

Consider now the projection of the curve on the rectifying plane. Taking formulas (3.23a) and (3.23c) and again limiting ourselves to the principal terms we obtain

$$x=\Delta l, \quad z=\frac{1}{6}k\varkappa\Delta l^3$$

Eliminating Δl from these relations we arrive at the equation of a cubic parabola:

$$z=\frac{1}{6}k\varkappa x^3 \tag{3.24}$$

The sign of the coefficient in x^3 coincides here with that of the torsion (because the curvature is always positive). The corresponding parabolas are shown in Fig. 3.8 for $\varkappa>0$ and $\varkappa<0$. The signs of the coordinates x and y being determined, for small values of Δl, by the signs of the corresponding principal terms, it follows from formula (3.24) that:

1. In the vicinity of a point at which the torsion is different from zero the curve lies on both sides of the osculating plane.

2. The greater the absolute value of the torsion, the greater the rate at which the curve is turning out of the osculating plane. If $\varkappa > 0$ the curve is turning out of the osculating plane, as l increases, in the direction of the vector $\boldsymbol{\beta}$, and in the opposite direction if otherwise.

Problems

1. Show that a curve whose curvature is identically equal to zero is a straight line.

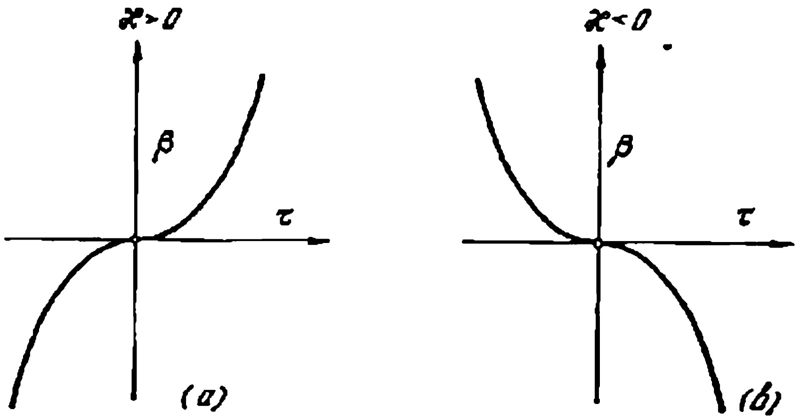


Fig. 3.8

2. Prove that a curve whose torsion is identically equal to zero is a plane curve, that is lies entirely in a fixed plane.

Solutions

1. If $k \equiv 0$ we have $\ddot{\mathbf{r}} \equiv 0$, i.e. $\dot{\mathbf{r}} = \mathbf{e} = \text{const}$ which implies $\mathbf{r} = \mathbf{r}_0 + l\mathbf{e}_0$, the last relation being an equation of a straight line.

2. If $\varkappa \equiv 0$, the third Frenet-Serret formula indicates that $\dot{\boldsymbol{\beta}} \equiv 0$, i.e. $\boldsymbol{\beta} = \boldsymbol{\beta}_0 = \text{const}$. The vectors $\dot{\mathbf{r}}$ and $\boldsymbol{\beta}_0$ being orthogonal, we have

$$(\boldsymbol{\beta}, \dot{\mathbf{r}}) = 0$$

Hence, since $\boldsymbol{\beta} = \boldsymbol{\beta}_0 = \text{const}$, we can write

$$\frac{d}{dl}(\boldsymbol{\beta}_0, \mathbf{r}) = 0$$

Consequently we arrive at the relation $(\boldsymbol{\beta}_0, \mathbf{r}) = \text{const}$ which is an equation of a plane.

7. **Curvature of a Plane Curve.** Consider a curve lying in a fixed plane. Introducing Cartesian coordinates x, y in the plane we can write the equation of the curve in the form

$$x = x(t), \quad y = y(t), \quad z \equiv 0 \tag{3.25}$$

Computing the curvature of the curve by means of formula (3.19) we find

$$k=\frac{|x'y''-x''y'|}{(x'^2+y'^2)^{3/2}} \tag{3.26}$$

But (e.g. see [8], Chapter 16, § 3) the curvature of a plane curve is usually defined with the absolute value sign removed, i.e. with the sign + or — attached to it. Then the expression for the curvature is written in the form

$$\frac{x'y''-x''y'}{(x'^2+y'^2)^{3/2}} \tag{3.27}$$

The matter is that in the case of the plane, unlike the case of the three-dimensional space, we can speak not only about the absolute value of the rate of turning of the tangent but also about its direction (i.e. the clockwise or the counter-clockwise direction). It is the

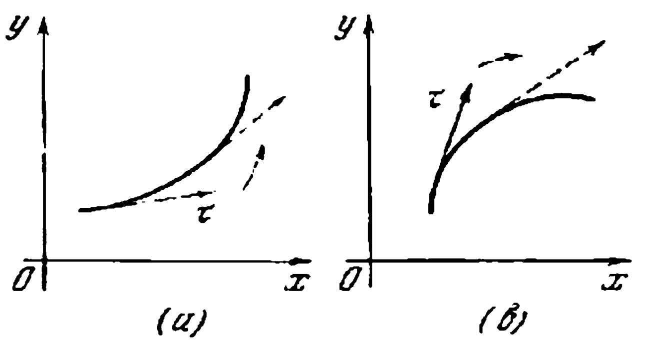


Fig. 3.9

direction of turning of the tangent that is indicated by the sign of quantity (3.27). A curve is said to be concave up if expression (3.27) is positive (see Fig. 3.9*a*) and concave down if otherwise (Fig. 3.9*b*).

8. **Intrinsic Equations of a Curve.** Formulas (3.13) and (3.15) make it possible to find the curvature and the torsion of a curve, given by an equation $\mathbf{r}=\mathbf{r}(l)$, as functions of l:

$$k=k(l), \quad \varkappa=\varkappa(l) \tag{3.28}$$

These relations connecting the curvature and the torsion of a curve with its arc length are known as **intrinsic (natural) equations** of the curve. We can now pose the question as to what extent intrinsic equations (3.28) determine the curve itself. It turns out that every curve is uniquely determined, to within its position in space, by its intrinsic equations.

Indeed, let us be given two curves γ and γ_1. Suppose it is possible to represent the curves parametrically by introducing the corresponding parameters l and l_1 (their arc lengths) in such a way that

their curvatures k, k_1 and torsions $\varkappa$, $\varkappa_1$ coincide at the point for which the parameters take equal values. In this case, for $l = l$ the relations

$$k(l) = k_1(l_1), \quad \varkappa(l) = \varkappa_1(l_1)$$

hold, and we say that the curves γ and γ_1 have the same intrinsi equations. Let us show that when this is so, one of the curves ca be made to coincide with the other if we move it in space, as a rigi body, in an appropriate way. Actually, apply a point A of th curve γ corresponding to a value l^0 of the parameter l to the poir A_1 of the curve γ_1 associated with the same value of the paramete $l_1 = l^0$. Further, turn the curve γ so that the unit vectors $\boldsymbol{\tau}$, $\mathbf{v}$, of its moving trihedron at the point A coincide with the correspon ing unit vectors $\boldsymbol{\tau}_1$, $\mathbf{v}_1$, $\boldsymbol{\beta}_1$ of the moving trihedron at the point A of the curve γ_1. Obviously, this can always be achieved. Then w have

$$\boldsymbol{\tau}^0 = \boldsymbol{\tau}_1^0, \quad \mathbf{v}^0 = \mathbf{v}_1^0, \quad \boldsymbol{\beta}^0 \quad \boldsymbol{\beta}_1^0 \tag{3.2}$$

where superscript "zero" indicates that the vectors are taken at th corresponding points determined by the common value $l^0 =$ of the parameters. Establishing the correspondence between th points M and M_1 of the curves γ and γ_1 for which $l = l_1$ we ca consider both curves to be represented parametrically with the hel of the same parameter l and thus regard $\boldsymbol{\tau}_1$, $\mathbf{v}_1$ and $\boldsymbol{\beta}_1$ as functio of l. Now take the scalar function

$$\sigma(l) = (\boldsymbol{\tau}, \boldsymbol{\tau}_1) + (\mathbf{v}, \mathbf{v}_1) + (\boldsymbol{\beta}, \boldsymbol{\beta}_1)$$

and find its derivative with respect to l. Taking advantage of t Frenet-Serret formulas we obtain

$$\frac{d\sigma}{dl} = k(\mathbf{v}, \boldsymbol{\tau}_1) + k(\mathbf{v}_1, \boldsymbol{\tau}) + (-k\boldsymbol{\tau} + \varkappa\boldsymbol{\beta}, \mathbf{v}_1) +$$

$$+ (\mathbf{v}, -k\boldsymbol{\tau}_1 + \varkappa\boldsymbol{\beta}_1) - \varkappa(\mathbf{v}, \boldsymbol{\beta}_1) - \varkappa(\mathbf{v}_1, \boldsymbol{\beta}) = 0$$

and thus σ does not in fact depend on l. Equalities (3.29) imply th the value of σ corresponding to $l = l_0$ is equal to three. Cons quently,

$$\sigma(l) \equiv 3$$

Each of the three summands entering into $\sigma(l)$ is a scalar produ of two unit vectors and hence cannot be greater than unity. T total sum being equal to three, each summand is exactly equal unity. But a scalar product of two unit vectors is equal to unit if and only if the vectors coincide. Therefore we have

$$\boldsymbol{\tau} \equiv \boldsymbol{\tau}_1, \quad \mathbf{v} \equiv \mathbf{v}_1, \quad \boldsymbol{\beta} \equiv \boldsymbol{\beta}_1$$

for all l, that is the moving trihedrons of the curves γ and γ_1 coinci not only at the initial point l_0 but also for all the values of t

parameter l. It follows that the curves themselves coincide because a curve can always be reconstructed from the vector $\boldsymbol{\tau} = \dot{\mathbf{r}}(l)$, namely

$$\mathbf{r}(l) = \mathbf{r}(l_0) + \int_{l_0}^{l} \boldsymbol{\tau}(\lambda)\, d\lambda$$

The converse is also true because if two curves differ only in their position in space, they obviously have the same intrinsic equations.

Now it seems natural to ask whether there exists a curve for which two arbitrarily chosen continuous functions

$$k(l) \quad (k(l) > 0) \quad \text{and} \quad \varkappa(l)$$

are, respectively, its curvature and torsion. It turns out that the answer to the question is affirmative but we shall not present the proof because this would involve some notions of the theory of differential equations that we are not going to study here.

9. Some Applications to Mechanics. Consider a material point moving along a trajectory. If $\mathbf{r}(t)$ is the radius vector of the point at moment of time t the equation of the trajectory is written in the form

$$\mathbf{r} = \mathbf{r}(t)$$

The derivative

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}(t)$$

is the velocity of motion of the point along the trajectory. Introducing the arc length as the parameter we can write

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{dl}\frac{dl}{dt} = \boldsymbol{\tau}\frac{dl}{dt}$$

Since $\boldsymbol{\tau}$ is a unit vector, we have

$$|\mathbf{v}| = \frac{dl}{dt}$$

and consequently the derivative $\frac{dl}{dt}$ expresses the absolute value of the velocity.

The second derivative

$$\mathbf{w} = \frac{d^2\mathbf{r}}{dt^2}$$

of the radius vector with respect to t is the acceleration of the point. It can be represented in the form

$$\mathbf{w} = \frac{d^2\mathbf{r}}{dl^2}\left(\frac{dl}{dt}\right)^2 + \boldsymbol{\tau}\frac{d^2l}{dt^2}$$

Applying the Frenet-Serret formulas we obtain

$$\mathbf{w} = \boldsymbol{\tau} \frac{d^2l}{dt^2} + \mathbf{v}k \left(\frac{dl}{dt}\right)^2$$

Thus, the acceleration is resolved into the sum of two components $\boldsymbol{\tau} \frac{d^2l}{dt^2}$ and $\mathbf{v}k \left(\frac{dl}{dt}\right)^2$. The former is in the direction of the tangent and is known as the **tangential acceleration** and the latter goes along the principal normal and is called the **normal (or centripetal) acceleration**. The tangential acceleration $\mathbf{w}_\tau = \boldsymbol{\tau} \frac{d^2l}{dt^2}$ can also be written as $\mathbf{w}_\tau = \boldsymbol{\tau} \frac{dv}{dt}$ where $v = \frac{dl}{dt}$ is the absolute value of the velocity which means that *the tangential acceleration is a measure of the rate of change of the absolute value of the velocity* $\mathbf{v}$. The formula $\mathbf{w}_\mathbf{v} = \mathbf{v}k \left(\frac{dl}{dt}\right)^2$ for the normal acceleration is well known from elementary courses of physics. Namely, when a point moves in a circle of radius R with a velocity $\mathbf{v}$ constant in its absolute value, $v = |\mathbf{v}|$,

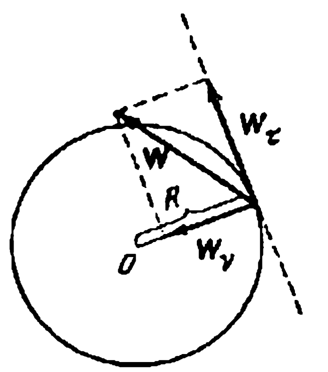


Fig. 3.10

the acceleration is always directed to the centre of the circle and its absolute value is equal to $\frac{1}{R} v^2$ where $\frac{1}{R}$ is nothing but the curva ture k of the circle. Hence, the resolution

$$\mathbf{w} = \mathbf{w}_\tau + \mathbf{w}_\mathbf{v} = \boldsymbol{\tau} \frac{d^2l}{dt^2} + \mathbf{v}k \left(\frac{dl}{dt}\right)^2$$

can be interpreted as follows: an arbitrary curvilinear motion is resolved, at a given moment of time, into an accelerated motion along the tangent (which results in the appearance of the term $\mathbf{w}_\tau$ in the acceleration) and a uniform motion in a circle of radius $R = \frac{1}{k}$ with the speed $v = \frac{dl}{dt}$ (which yields the term $\mathbf{w}_\mathbf{v}$ in the acceleration). Hence, the point simultaneously takes part in the two motions (see Fig. 3.10).

Problem. A mass point moves under the action of a central force, i.e. one whose line of action always passes through a fixed centre. Prove that the trajectory is a plane curve.

Solution. Take the centre as the origin of coordinates. Let

$$\mathbf{r} = \mathbf{r}\,(t)$$

be the equation of the trajectory of motion. The force acting upon the moving point is directed toward the centre. Consequently, according to Newton's second law, the acceleration, that is the vector $\mathbf{r}''(t)$, has the same direction. Therefore the vectors $\mathbf{r}$ and $\mathbf{r}''$ are collinear and hence the relation

$$(\mathbf{r},\ \mathbf{r}',\ \mathbf{r}'') = 0$$

holds at each point of the trajectory. Differentiating the triple scalar product with respect to t we obtain

$$\frac{d}{dt}(\mathbf{r},\ \mathbf{r}',\ \mathbf{r}'') = (\mathbf{r}',\ \mathbf{r}',\ \mathbf{r}'') + (\mathbf{r},\ \mathbf{r}'',\ \mathbf{r}'') + (\mathbf{r},\ \mathbf{r}',\ \mathbf{r}''') = 0$$

The terms $(\mathbf{r}',\ \mathbf{r}',\ \mathbf{r}'')$ and $(\mathbf{r},\ \mathbf{r}'',\ \mathbf{r}'')$ are equal to zero and consequently

$$(\mathbf{r},\ \mathbf{r}',\ \mathbf{r}''') = 0$$

The vectors $\mathbf{r}$ and $\mathbf{r}''$ being collinear, we thus have

$$(\mathbf{r}',\ \mathbf{r}'',\ \mathbf{r}''') = 0$$

for all t. It follows that $\varkappa \equiv 0$, the last relation indicating that the trajectory is a plane curve (see § 2, Sec. 6).

§ 3. PARAMETRIC EQUATIONS OF A SURFACE

1. The Concept of a Surface. The present and subsequent sections of this chapter are devoted to the application of mathematical analysis to studying surfaces.

The concept of surface, although clear enough from an intuitive point of view, can be defined with various degrees of generality. In mathematical analysis we often consider surfaces represented by an equation of the form

$$z = f\,(x,\ y)$$

where $f\,(x,\ y)$ is a continuous function defined in a domain G. A wider class of surfaces is described by equations of the form

$$F\,(x,\ y,\ z) = 0$$

For such an equation to determine a surface, as we intuitively understand it, it is necessary that the function $F\,(x,\ y,\ z)$ satisfy some additional requirements.

The definition of a surface as a set of points whose coordinates satisfy an equation of the form $z = f\,(x,\ y)$ or $F\,(x,\ y,\ z) = 0$ is sometimes inconvenient because it is closely related to the specific

choice of the coordinate system. Therefore we shall give a definition of the concept of surface without involving a coordinate system. First of all we introduce an important notion of a *simply connected domain*.

Let G be a domain in the plane. We say that the domain G is simply connected if the following condition is satisfied: every closed contour L lying inside the domain bounds a (finite) part of the plane entirely contained in G.

In other words, a simply connected domain is one without "holes". Every closed contour lying inside such a domain can be continuously contracted to a point without falling outside the domain.

A domain which is not simply connected is referred to as a multiply connected domain.

Examples of a simply connected domain are the circle, the triangle, the square and so on. An annulus, that is a part of the plane bounded by two concentric circles, is an example of a multiply connected

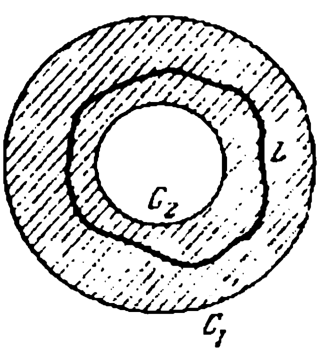


Fig. 3.11

domain. Indeed, the part of the plane bounded by the contour L shown in Fig. 3.11 is by no means a part of the annulus lying between the circles C_1 and C_2.

By a *simple surface* we shall understand a set of points in a three-dimensional space which is representable as an image of a bounded closed simply connected domain under a one-to-one bicontinuous (i.e. continuous in both directions) mapping. Further, the term *surface* will be applied to every union of a finite number of simple surfaces. This also includes the case of self-intersecting surfaces. For instance, we can consider such geometric configurations as the one shown in Fig. 3.12.

If $f(x, y)$ is a continuous function defined in a bounded closed domain G the equation

$$z = f(x, y)$$

determines a simple surface. In fact, the mapping

$$(x, y) \leftrightarrow (x, y, f(x, y))$$

specifies a one-to-one correspondence, continuous in both directions, between the points (x, y) of the domain G and the points (x, y, z) whose coordinates satisfy the equation $z = f(x, y)$ (check it up).

Practically, in what follows we shall restrict our consideration to surfaces representable as a union of a finite number of simple surfaces determined by equations of the form $z = f(x, y)$. Besides the condition of continuity, we shall usually impose some requirements specifying the smoothness of the corresponding functions f

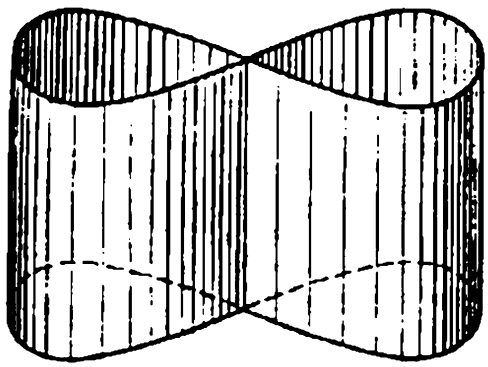


Fig. 3.12

(the existence and the continuity of its first or first and second partial derivatives). Such conditions will be explicitly stipulated when necessary.

2. Parametrization of a Surface. Although in mathematical analysis we very often deal with surfaces defined by equations of the form $z = f(x, y)$ or $F(x, y, z) = 0$ it is sometimes more convenient to determine a surface by means of *parametric equations.* To write down a parametric representation of a surface we first introduce the notion of *coordinates on a surface.*

Suppose there is a *one-parameter family of curves** lying on a surface Σ. We shall say that the family is regular if, for every given point of the surface, there is one and only one curve belonging to

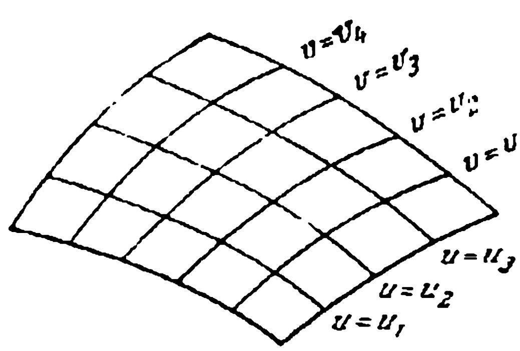


Fig. 3.13

the family which passes through the point. If there are two regular families on a surface such that each curve of one family has a single common point (point of intersection) with each curve of the other family and the curves are not tangent to each other at the points of intersection we say that there is a system of parametric (coordi-

* This means that each curve of the family is characterized by a certain value of a single parameter.

nate) curves on the surface. Let the curves of one family be determined by the values of a parameter u (we call them *u-curves*) and the curves of the other family by the values of a parameter v (*v-curves*; see Fig. 3.13). By the hypothesis, for every given point of the surface, there is a single curve of one family and a single curve of the other family passing through the point, and therefore the position of each point on the surface is uniquely determined by certain values u_0 and v_0 of the parameters u and v corresponding to the curves. The parameters u and v whose values specify the curves are called **(curvilinear) coordinates** on the surface.

Note. In § 6 of Chapter 1 we introduced curvilinear coordinates in a plane region (domain). Here we have repeated the construction but applied it to a curvilinear surface in space. The introduction of coordinates on a surface is obviously equivalent to the specification of a one-to-one continuous mapping of the surface onto a part of the plane where the Cartesian coordinates u and v have been introduced. The parametric curves forming the system of coordinate curves on the surface are the images of the straight lines parallel to the coordinate axes in the u, v-plane.

Examples

1. A **torus (anchor ring)** is a surface generated by the rotation, in space, of a circle about an axis in its plane but not cutting the circle. The position of a point on the circle can be determined by an angle φ ($0 \leqslant \varphi < 2\pi$) reckoned from an initial point. The position

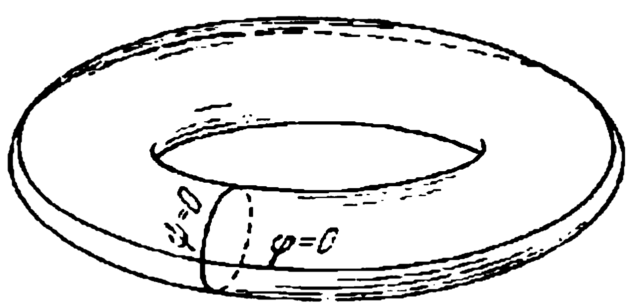


Fig. 3.14

of the circle itself can be specified by the angle of turn ψ reckoned from its initial position. Thus, the position of a current point on the torus is determined by the two angles φ and ψ independently varying within the limits from 0 to 2π. The curves $\varphi = 0$ and $\psi = 0$ of the corresponding families of parametric curves are depicted in Fig. 3.14.

2. Let a surface be represented by an equation $z = f(x, y)$. In other words, let there be a one-to-one correspondence between its points and the points of its projection on an x, y-plane. The curves whose projections are the straight lines $x = \text{const}$ and $y = \text{const}$ form the corresponding families of coordinate curves on the surface $z = f(x, y)$ (see Fig. 3.15).

It is clear that there are various families of coordinate curves that can be constructed on the same surface.

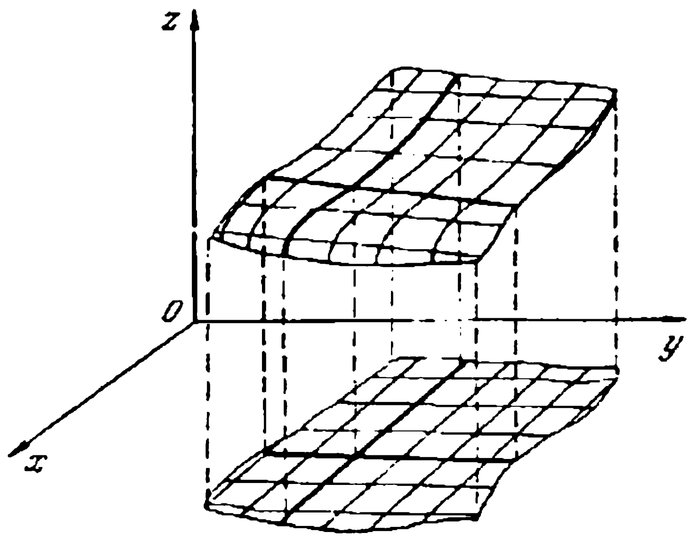


Fig. 3.15

3. Parametric Equations of a Surface. If some coordinates u and v are introduced in a certain way on a surface Σ we can write a so-called **parametric representation of the surface** corresponding to the parameters u and v. Each point of the surface can be determined by certain values of the parameters u and v but at the same time it can be specified by its Cartesian coordinates. Consequently, the Cartesian coordinates of the points of a surface, on which some curvilinear coordinates u and v have been introduced, are functions of the coordinates in the u, v-plane:

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v) \tag{3.30}$$

The three scalar equations can be replaced by a single vector equation:

$$\mathbf{r} = \mathbf{r}(u, v) \tag{3.30$'$}$$

where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Equations of form (3.30) or (3.30′) will be referred to as **parametric equations of the surface.**

Note 1. When we write a parametric equation of a curve the coordinates x, y, z are functions of a single parameter. An equation $\mathbf{r} = \mathbf{r}(u, v)$ of a surface naturally involves two parameters since a surface is a two-dimensional geometric configuration.

Note 2. An equation $z = f(x, y)$ can be considered a special case of a parametric equation because, if we choose x and y as parameters, we can write

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + f(x, y)\,\mathbf{k}$$

Exercise. Write the parametric equations of a torus (see example 1 in Sec. 2) in coordinates φ and ψ.

As a rule, in what follows we shall consider surfaces represented by parametric equations. The function $\mathbf{r}(u, v)$ will be supposed

to be continuous together with its partial derivatives of the first order with respect to u and v. In § 5 and in the subsequent sections of the present chapter we shall additionally impose the condition of existence and continuity of the partial derivatives of the second order.

4. Curves on a Surface. Consider a curve on a surface determined by equation (3.30′). If a parameter t is introduced on the curve then, to each value of t, there corresponds a point of the surface, i.e. certain values of u and v. Thus, the coordinates u and v become functions of the parameter t when considered along the curve:

$$u = u(t), \quad v = v(t)$$

These equations are the *equations of the curve on the surface*. Substituting them into the equation of the surface we arrive at a *parametric equation of the curve on the surface*:

$$\mathbf{r} = \mathbf{r}(u(t), v(t)) \tag{3.31}$$

Conversely, substituting arbitrary functions of a single variable t for the independent variables u and v into equation (3.30′) of the surface we obtain the equation of a curve lying on the surface.

Let us consider the tangent to curve (3.31). Its direction is determined by the vector

$$\frac{d\mathbf{r}}{dt} = \frac{\partial \mathbf{r}}{\partial u}\frac{du}{dt} + \frac{\partial \mathbf{r}}{\partial v}\frac{dv}{dt}$$

which is a linear combination of the vectors $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ called **base vectors** (for the curvilinear coordinates in question). At each point, the vectors are tangent to the coordinate curves passing through the point. We denote them, for brevity, as $\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u}$ and $\mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v}$.

5. Tangent Plane. Consider all the possible curves lying on a surface and passing through a given point M and the tangent vectors to the curves at the point (see Fig. 3.16). Each vector can be expressed linearly in terms of the vectors $\mathbf{r}_u$ and $\mathbf{r}_v$, and hence it lies in the plane determined by the vectors. This plane is said to be the **tangent plane to the surface at the point** M. Let us form the equation of the tangent plane. The vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ lying in the tangent plane, the vector $\mathbf{N} = [\mathbf{r}_u, \mathbf{r}_v]$ is orthogonal to the plane, and hence the sought-for equation of the plane is

$$(\boldsymbol{\rho} - \mathbf{r}, \mathbf{N}) = 0 \tag{3.32}$$

where $\mathbf{r}$ is the radius vector of the point of tangency and $\boldsymbol{\rho}$ is the radius vector of the moving point in the tangent plane.*

---

* Here and henceforward we exclude from our consideration the points at which $[\mathbf{r}_u, \mathbf{r}_v] = 0$

Now suppose that a surface is represented by an equation $z = f(x, y)$ which can be written in vector form as

$$\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}f(x, y)$$

Let us write the equation of the tangent plane to such a surface. We have

$$\mathbf{r}_x = \mathbf{i} + \mathbf{k}f'_x, \quad \mathbf{r}_y = \mathbf{j} + \mathbf{k}f'_y$$

and, consequently,

$$\mathbf{N} = [\mathbf{r}_x, \mathbf{r}_y] = -\mathbf{i}f'_x - \mathbf{j}f'_y + \mathbf{k} \tag{3.33}$$

Substituting the vector $\mathbf{i}(x - x_0) + \mathbf{j}(y - y_0) + \mathbf{k}(z - z_0)$ for $\boldsymbol{\rho} - \mathbf{r}$ into equation (3.32) of the tangent plane, and expression (3.33)

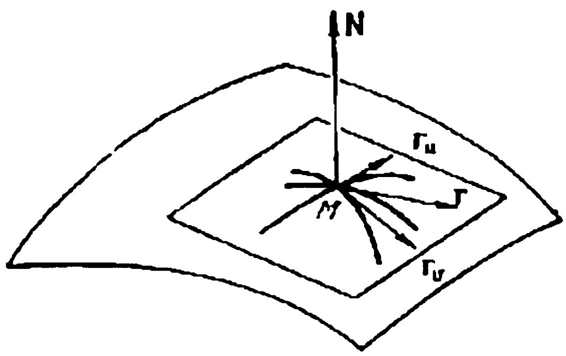


Fig. 3.16

for the vector $\mathbf{N}$, we obtain the equation of the tangent plane to the surface $z = f(x, y)$ at the point (x_0, y_0, z_0):

$$z - z_0 = f'_x(x - x_0) + f'_y(y - y_0) \tag{3.34}$$

where the values of the partial derivatives f'_x and f'_y are taken at the point (x_0, y_0) (the projection of the point of tangency (x_0, y_0, z_0) on the x, y-plane).

If a surface is determined by an equation $F(x, y, z) = 0$, defining z as an implicit function of x and y, we can write

$$\frac{\partial z}{\partial x} = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial z}}, \qquad \frac{\partial z}{\partial y} = -\frac{\dfrac{\partial F}{\partial y}}{\dfrac{\partial F}{\partial z}}$$

Substituting these expressions for f'_x and f'_y into equation (3.33) we derive the equation of the tangent plane to the surface $F(x, y, z) = 0$ at an arbitrary point (x_0, y_0, z_0) (where $F(x_0, y_0, z_0) = 0$):

$$(x - x_0)F'_x + (y - y_0)F'_y + (z - z_0)F'_z = 0$$

The values F'_x, F'_y and F'_z entering into the formula are taken at the point of tangency (x_0, y_0, z_0).

6. Normal to a Surface. Consider a surface $\mathbf{r} = \mathbf{r}(u, v)$. Let us find the direction cosines of the vector

$$\mathbf{N} = [\mathbf{r}_u, \mathbf{r}_v]$$

perpendicular to the tangent plane of the surface $\mathbf{r} = \mathbf{r}(u, v)$. It is called **the vector of the normal (normal line) to the surface**, which is a straight line passing through the point $\mathbf{r} = \mathbf{r}(u, v)$ of the surface perpendicularly to its tangent plane. We have

$$\mathbf{r}_u = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right) \quad \text{and} \quad \mathbf{r}_v = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right)$$

and consequently the projections of the vector $[\mathbf{r}_u, \mathbf{r}_v]$ are

$$A = \begin{vmatrix} \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{vmatrix}, \quad B = \begin{vmatrix} \frac{\partial z}{\partial u} & \frac{\partial x}{\partial u} \\ \frac{\partial z}{\partial v} & \frac{\partial x}{\partial v} \end{vmatrix}, \quad C = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} \tag{3.35}$$

Therefore its direction cosines are respectively equal to the expressions

$$\cos(\mathbf{N}, x) = \frac{A}{\sqrt{A^2 + B^2 + C^2}}, \quad \cos(\mathbf{N}, y) = \frac{B}{\sqrt{A^2 + B^2 + C^2}},$$
$$\cos(\mathbf{N}, z) = \frac{C}{\sqrt{A^2 + B^2 + C^2}}$$

In particular, if a surface is given by an equation

$$z = f(x, y)$$

or, in vector form, by the corresponding equation

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + f(x, y)\,\mathbf{k}$$

we have

$$A = \begin{vmatrix} 0 & f'_x \\ 1 & f'_y \end{vmatrix} = -f'_x, \quad B = \begin{vmatrix} f'_x & 1 \\ f'_y & 0 \end{vmatrix} = -f'_y, \quad C = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

Thus, in this case the formulas for the direction cosines are

$$\cos(\mathbf{N}, x) = \frac{-f'_x}{\sqrt{1 + f'^2_x + f'^2_y}}, \quad \cos(\mathbf{N}, y) = \frac{-f'_y}{\sqrt{1 + f'^2_x + f'^2_y}},$$
$$\cos(\mathbf{N}, z) = \frac{1}{\sqrt{1 + f'^2_x + f'^2_y}} \tag{3.36}$$

7. Coordinate Systems in Tangent Planes. Consider a surface Σ having the tangent plane at each point M. It is sometimes convenient to think of a surface as being covered by the "scales" formed of tangent planes. Thus, the surface is interpreted as a curvilinear manifold which is the carrier of its tangent planes, the latter being

the carried linear (plane) manifolds. Such an approach will be of use for studying the subject matter of this and the subsequent sections of the present chapter.

Choose a pair of noncollinear vectors e_1 and e_2 in each tangent plane and consider them the base vectors of a coordinate system in the plane. These base vectors can be taken at pleasure at each point. But if the surface is defined by parametric equations with parameters u and v we can construct a specific pair of base vectors generated in a natural manner at each point by the parametric representation of the surface, namely the vectors $\mathbf{e}_1 = \frac{\partial \mathbf{r}}{\partial u}$ and $\mathbf{e}_2 = \frac{\partial \mathbf{r}}{\partial v}$. If we fix a value $v = v_0$ of the parameter v and make

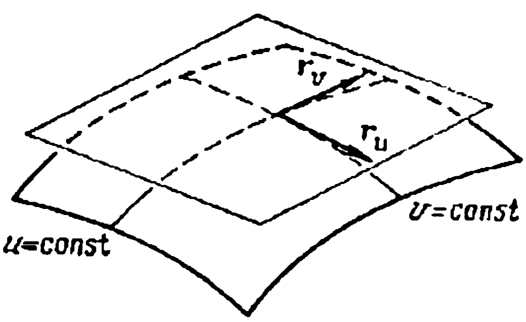


Fig. 3.17

the parameter u vary the radius vector $\mathbf{r}(u, v_0)$ describes the coordinate curve $v = v_0 =$ const on the surface (Fig. 3.17). The tangent vector to the curve, i.e. $\frac{\partial \mathbf{r}}{\partial u}(u, v_0)$, lies in the tangent plane to the surface (see Sec. 4). Similarly, the vector $\frac{\partial \mathbf{r}}{\partial u}$ also lies in the tangent plane to the surface. As before, we suppose that only a single curve of each of the families $u =$ const and $v =$ const passes through every point. Therefore we have a uniquely defined pair of base vectors $\mathbf{r}_u$, $\mathbf{r}_v$ in each tangent plane. If the vectors are different from zero they are noncollinear since, according to the hypothesis, the curves $u =$ const and $v =$ const are at no point tangent to each other. Hence, they may turn out to be collinear only when one of them or both vanish. In what follows we shall suppose that the parametric representation is such that $\mathbf{r}_u \neq 0$ and $\mathbf{r}_v \neq 0$ on the piece of the surface we are dealing with.

Thus, every parametric representation of a surface with parameters u and v generates a uniquely determined pair of base vectors $\mathbf{e}_1 = \mathbf{r}_u$, $\mathbf{e}_2 = \mathbf{r}_v$, i.e. an affine coordinate system, in each tangent plane to the surface.

If some other parameters $\tilde{u}$ and $\tilde{v}$ are chosen instead of u and v we obtain another set of coordinate systems determined by the base vectors $\tilde{\mathbf{e}}_1 = \mathbf{r}_{\tilde{u}}$ and $\tilde{\mathbf{e}}_2 = \mathbf{r}_{\tilde{v}}$ in the tangent planes. The transition from one parametric representation to another generates an affine transformation of coordinate systems.

Actually, let

$$u = u(\tilde{u}, \tilde{v}), \quad v = v(\tilde{u}, \tilde{v})$$

be the expressions of the parameters u, v in terms of $\tilde{u}$, $\tilde{v}$. According to the rule of differentiation of a composite vector function we obtain

$$\left.\begin{aligned} \mathbf{r}_{\tilde{u}} &= \mathbf{r}_u \frac{\partial u}{\partial \tilde{u}} + \mathbf{r}_v \frac{\partial v}{\partial \tilde{u}} \\ \mathbf{r}_{\tilde{v}} &= \mathbf{r}_u \frac{\partial u}{\partial \tilde{v}} + \mathbf{r}_v \frac{\partial v}{\partial \tilde{v}} \end{aligned}\right\} \tag{3.37}$$

Consequently, the new base vectors $\tilde{\mathbf{e}}_1$ and $\tilde{\mathbf{e}}_2$ are expressed in terms of $\mathbf{e}_1$ and $\mathbf{e}_2$ by the formulas

$$\left.\begin{aligned} \tilde{\mathbf{e}}_1 &= \frac{\partial u}{\partial \tilde{u}} \mathbf{e}_1 + \frac{\partial v}{\partial \tilde{u}} \mathbf{e}_2 \\ \tilde{\mathbf{e}}_2 &= \frac{\partial u}{\partial \tilde{v}} \mathbf{e}_1 + \frac{\partial v}{\partial \tilde{v}} \mathbf{e}_2 \end{aligned}\right\} \tag{3.37'}$$

The formulas expressing the base vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ as linear combinations of $\tilde{\mathbf{e}}_1$ and $\tilde{\mathbf{e}}_2$ are similar to (3.37′).

§ 4. DETERMINING LENGTHS, ANGLES AND AREAS ON A CURVILINEAR SURFACE. FIRST FUNDAMENTAL QUADRATIC FORM OF A SURFACE

There are many physical, technical and geometric problems involving the computation of arc lengths for curves lying on a surface, angles between such curves and areas of various parts of the surface. Here we shall discuss these questions. The key idea of all the considerations given in § 4 is essentially based on replacing an infinitesimal element of a smooth surface by the corresponding element of its tangent plane. It is therefore expedient to begin with some formulas and notions related to determining lengths, angles and areas in the plane.

1. Affine Coordinate System in the Plane. Consider a plane and a pair of noncollinear base vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ lying in it. Every vector in the plane is expressible in the form

$$\mathbf{r} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2$$

Let us find the square of the length of the vector $\mathbf{r}$. We have

$$r^2 = (\mathbf{r}, \mathbf{r}) = \xi_1^2 (\mathbf{e}_1, \mathbf{e}_1) + 2\xi_1\xi_2 (\mathbf{e}_1, \mathbf{e}_2) + \xi_2^2 (\mathbf{e}_2, \mathbf{e}_2)$$

Introducing the notation

$$g_{11} = (\mathbf{e}_1, \mathbf{e}_1), \quad g_{12} = (\mathbf{e}_1, \mathbf{e}_2), \quad g_{22} = (\mathbf{e}_2, \mathbf{e}_2)$$

we rewrite the last relation in the form

$$r^2 = g_{11}\xi_1^2 + 2g_{12}\xi_1\xi_2 + g_{22}\xi_2^2 \tag{3.38}$$

The quantities g_{11}, g_{12} and g_{22} (referred to as **metric coefficients**) are specified by the choice of the base vectors $\mathbf{e}_1$ and $\mathbf{e}_2$. It can be easily shown that the lengths of the vectors lying in the plane, the angles between the vectors and the areas of the parallelograms constructed on the vectors are expressible in terms of these quantities (and, of course, in terms of the coordinates of the corresponding vectors). Indeed, the expression for the length r of the vector $\mathbf{r}$ is obtained from formula (3.38). Further, if

$$\mathbf{r} = \xi_1\mathbf{e}_1 + \xi_2\mathbf{e}_2 \quad \text{and} \quad \boldsymbol{\rho} = \eta_1\mathbf{e}_1 + \eta_2\mathbf{e}_2$$

we have

$$(\mathbf{r}, \boldsymbol{\rho}) = g_{11}\xi_1\eta_1 + g_{12}\xi_1\eta_2 + g_{12}\xi_2\eta_1 + g_{22}\xi_2\eta_2$$

Now applying the formula

$$\cos(\mathbf{r}, \boldsymbol{\rho}) = \frac{(\mathbf{r}, \boldsymbol{\rho})}{|\mathbf{r}||\boldsymbol{\rho}|}$$

we can express the angle between the vectors $\mathbf{r}$ and $\boldsymbol{\rho}$ in terms of their coordinates and the coefficients g_{ik}.

Finally, let us find the area S of the parallelogram constructed on the vectors $\mathbf{r}$ and $\boldsymbol{\rho}$. As is well known,

$$S = |[\mathbf{r}, \boldsymbol{\rho}]|$$

and therefore

$$S = |[\xi_1\mathbf{e}_1 + \xi_2\mathbf{e}_2, \eta_1\mathbf{e}_1 + \eta_2\mathbf{e}_2]| = |\xi_1\eta_2 - \xi_2\eta_1| \, |[\mathbf{e}_1, \mathbf{e}_2]|$$

But

$$|[\mathbf{e}_1, \mathbf{e}_2]| = |\mathbf{e}_1||\mathbf{e}_2| \sin(\mathbf{e}_1, \mathbf{e}_2) = |\mathbf{e}_1||\mathbf{e}_2| \sqrt{1 - \cos^2(\mathbf{e}_1, \mathbf{e}_2)} =$$

$$= \sqrt{e_1^2 e_2^2 - (\mathbf{e}_1, \mathbf{e}_2)^2} = \sqrt{g_{11}g_{22} - g_{12}^2}$$

Consequently,

$$S = \sqrt{g_{11}g_{22} - g_{12}^2}\, |\xi_1\eta_2 - \xi_2\eta_1|$$

Thus, we can really find the lengths, the angles and the areas on the plane when the quantities g_{11}, g_{12} and g_{22} are known.*

---

* We sometimes use the notation

$$E = (\mathbf{e}_1, \mathbf{e}_1), \quad F = (\mathbf{e}_1, \mathbf{e}_2), \quad G = (\mathbf{e}_2, \mathbf{e}_2)$$

Putting $g_{21} = g_{12}$ we can also write the quantities g_{ik} $(i, k = 1, 2)$ in the form of a matrix

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

2. Arc Length of a Curve on a Surface. First Fundamental Quadratic Form. Let us be given a surface

$$\mathbf{r} = \mathbf{r}(u, v)$$

Compute the arc length of a curve lying on the surface. Taking the arc length of the curve as the parameter we can write its equation in the form

$$\mathbf{r} = \mathbf{r}(u(l), v(l))$$

The vector $\frac{d\mathbf{r}}{dl}$ being of unit length, we have

$$dl^2 = d\mathbf{r}^2$$

But

$$d\mathbf{r} = \mathbf{r}_u\, du + \mathbf{r}_v\, dv$$

and consequently

$$dl^2 = r_u^2\, du^2 + 2(\mathbf{r}_u, \mathbf{r}_v)\, du\, dv + r_v^2\, dv^2$$

Making use of the notation

$$g_{11} = r_u^2, \quad g_{12} = (\mathbf{r}_u, \mathbf{r}_v), \quad g_{22} = r_v^2$$

we obtain

$$dl^2 = g_{11}\, du^2 + 2g_{12}\, du\, dv + g_{22}\, dv^2 \tag{3.39}$$

This expression is a quadratic form (in the variables du and dv) which is apparently *positive definite**. It is called the **first fundamental quadratic form of the surface** $\mathbf{r} = \mathbf{r}(u, v)$. The coefficients g_{11}, g_{12} and g_{22} of the form are obviously the ones corresponding to the base vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ in the tangent plane to the surface at the point in question. The coefficients may vary as the point moves on the surface. Besides, they are of course dependent on the specific choice of the parametric representation of the surface.

The first fundamental quadratic form of a surface provides the expression for the length of an infinitesimal arc. The length of a finite curve lying on the surface is obtained from it by integration. More precisely, if a curve on a surface is given by equations

$$u = u(t), \quad v = v(t), \quad t_1 \leqslant t \leqslant t_2$$

its length is equal to

$$l = \int_{t_1}^{t_2} \sqrt{g_{11}\left(\frac{du}{dt}\right)^2 + 2g_{12}\frac{du}{dt}\frac{dv}{dt} + g_{22}\left(\frac{dv}{dt}\right)^2}\, dt$$

(the quantities g_{11}, g_{12} and g_{22} become functions of the parameter t when the current point moves along the curve).

---

* A quadratic form $\sum_{i,k=1}^{n} a_{ik}\xi_i\xi_k$ is said to be positive definite if $\sum_{i,k=1}^{n} a_{ik}\xi_i\xi_k > 0$ for all $\xi_1, \xi_2, \ldots, \xi_n$ except $\xi_1 = \xi_2 = \ldots = \xi_n = 0$.

Examples

1. Let, in the plane, there be given a coordinate system determined by two mutually orthogonal unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$. If $\mathbf{r}_0$ is the radius vector of the origin of the coordinates the radius vector of an arbitrary point is equal to

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{e}_1 u + \mathbf{e}_2 v$$

We have thus obtained a parametric representation of the plane, the parameters being the Cartesian coordinates u and v.

In this case we have

$$\mathbf{r}_u = \mathbf{e}_1, \quad \mathbf{r}_v = \mathbf{e}_2, \quad g_{11} = 1, \quad g_{12} = 0, \quad g_{22} = 1$$

and consequently the first fundamental quadratic form of the plane represented parametrically by means of its Cartesian coordinates is written as

$$dl^2 = du^2 + dv^2$$

In this example the tangent plane coincides with the surface (which is also a plane) at all the points, and the pair of base vectors generated in each tangent plane by the parametric representation coincides (to within a parallel translation) with the base vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ chosen in the plane.

2. Introduce polar coordinates ρ and φ in the plane. Then the radius vector of an arbitrary point can be written in the form

$$\mathbf{r} = \mathbf{r}_0 + \rho\,(\mathbf{e}_1 \cos\varphi + \mathbf{e}_2 \sin\varphi)$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are again mutually orthogonal unit vectors. This is the equation of the plane represented parametrically by means of polar coordinates. Here we have

$$\mathbf{r}_\rho = \mathbf{e}_1 \cos\varphi + \mathbf{e}_2 \sin\varphi, \quad \mathbf{r}_\varphi = \rho\,(-\mathbf{e}_1 \sin\varphi + \mathbf{e}_2 \cos\varphi)$$

and consequently

$$g_{11} = (\mathbf{r}_\rho, \mathbf{r}_\rho) = 1, \quad g_{12} = (\mathbf{r}_\rho, \mathbf{r}_\varphi) = 0,$$
$$g_{22} = (\mathbf{r}_\varphi, \mathbf{r}_\varphi) = \rho^2, \quad dl^2 = d\rho^2 + \rho^2\, d\varphi^2$$

3. Consider a sphere of radius a and take the longitude φ and the latitude θ as the parameters on it* (see Fig. 3.18). The equation of the sphere in the coordinates φ and θ is of the form

$$\mathbf{r} = \mathbf{r}_0 + a\{(\mathbf{i} \cos\varphi + \mathbf{j} \sin\varphi) \cos\theta + \mathbf{k} \sin\theta\}$$

(check it up). It follows that

$$\mathbf{r}_\theta = -a\,(\mathbf{i} \cos\varphi + \mathbf{j} \sin\varphi) \sin\theta + a\mathbf{k} \cos\theta$$
$$\mathbf{r}_\varphi = a\,(-\mathbf{i} \sin\varphi + \mathbf{j} \cos\varphi) \cos\theta$$

---

* Let the latitude θ be reckoned from the equator, i.e. $-\frac{\pi}{2} \leqslant \theta \leqslant \frac{\pi}{2}$.

and hence here we have

$$dl^2 = a^2 \,(d\theta^2 + \cos^2 \theta \, d\varphi^2)$$

4. If a surface is given by an expression of z as an explicit function

$$z = f\,(x, y)$$

that is

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + f\,(x, y)\,\mathbf{k}$$

we can write the relations

$$\mathbf{r}_x = \mathbf{i} + \mathbf{k} f'_x, \quad \mathbf{r}_y = \mathbf{j} + \mathbf{k} f'_y$$

and, consequently,

$$dl^2 = (1 + f_x'^2)\, dx^2 + 2 f'_x f'_y \, dx\, dy + (1 + f_y'^2)\, dy^2$$

Exercise. Write the first fundamental quadratic form for a torus in the coordinates φ and ψ (see the exercise in § 3, Sec. 3).

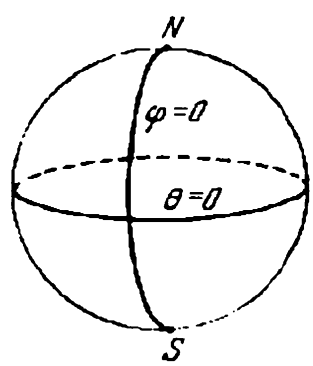


Fig. 3.18

3. Angle Between Two Curves. The angle between two intersecting curves is, by definition, the one formed by their tangent lines at the point of intersection. Suppose two curves lying on a surface have a point in common. Let du and dv be the differentials of the coordinates corresponding to a displacement from the point of intersection along one curve, and δu and δv be the differentials of the coordinates corresponding to a displacement along the other curve. The displacement vectors can be written as

$$d\mathbf{r} = \mathbf{r}_u \, du + \mathbf{r}_v \, dv, \quad \delta \mathbf{r} = \mathbf{r}_u \, \delta u = \mathbf{r}_v \, \delta v$$

The angle φ between them is determined by the formula

$$\cos \varphi = \frac{(d\mathbf{r}, \delta\mathbf{r})}{|d\mathbf{r}|\,|\delta\mathbf{r}|}$$

In particular, the angle ω between the coordinate curves, that is between the vectors $\mathbf{r}_u$ and $\mathbf{r}_v$, is determined by the formula

$$\cos \omega = \frac{g_{12}}{\sqrt{g_{11} g_{22}}}$$

If $g_{12} \equiv 0$ the coordinate curves on the surface intersect at a right angle. In this case we have so-called **curvilinear orthogonal coordinates**. The first fundamental quadratic form is expressed, in orthogonal coordinates, by the formula

$$dl^2 = g_{11}\, du^2 + g_{22}\, dv^2$$

4. Definition of Area of a Surface. The Schwarz Example. We now proceed to study the notion of area for a curvilinear surface. Before discussing the ways of computing the area we must give the

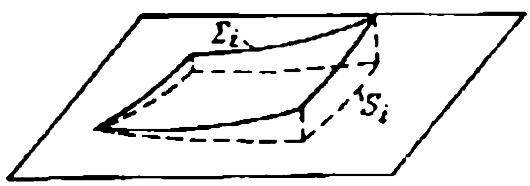


Fig. 3.19

definition of the notion. It is introduced as follows. Let Σ be a smooth surface bounded by a piecewise smooth contour L. Break up the surface into "elements" Σ_i, $i = 1, \ldots, N$, by means of a finite number of piecewise smooth curves lying on the surface, and choose a point M_i in each part. Next draw the tangent plane to the surface Σ through each point M_i and project the elements Σ_i on the corresponding tangent planes. We thus obtain squarable figures S_i in the tangent planes (Fig. 3.19).

Definition. *The **area of the surface** Σ is the limit (provided it exists) of the sum of the areas of the projections extended over all the elements Σ_i of the partition $\{\Sigma_i\}$ when the maximal of the diameters of the elements tends to zero. A surface for which the limit exists is said to be **squarable (rectifiable)**.*

One may think that it would be more natural to define the area of an arbitrary surface Σ as the limit to which the areas of the surfaces of the polyhedrons inscribed in Σ tend, on condition that the maximal of the diameters of their faces tends to zero (by analogy with the arc length of a curve which is the limit of the lengths of the broken lines inscribed in the curve). But as early as the 19th century it was found that such a definition was inconsistent. Consider the following example of Schwarz*.

Let us inscribe a polyhedron in a cylinder of radius R and altitude H in the following way. Divide the cylinder into m equal parts by means of horizontal planes, the altitude of each part being $\frac{H}{m}$ (see Fig. 3.20). Break up each of the $m+1$ circles appearing in the sections (including the upper and the lower bases of the original cylinder) into n equal parts so that the points of division of each

* Schwarz, Hermann Amandus (1843-1921), a German mathematician.

circle are placed above the midpoints of the arcs of the adjacent circle. Now take two neighbouring points lying on a circle and the midpoint of the corresponding arc of the nearest circle lying above or below the circle and construct a triangle with vertices at the three points. The union of these triangles is a polyhedral surface shown in Fig. 3.21.

If now n and m are infinitely increased the sizes of all the triangles (which are the faces of the polyhedral surface inscribed in the cylinder) tend to zero. But the total area of all the triangles by far not

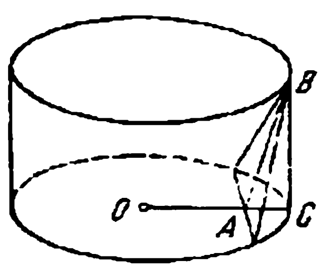


Fig. 3.20

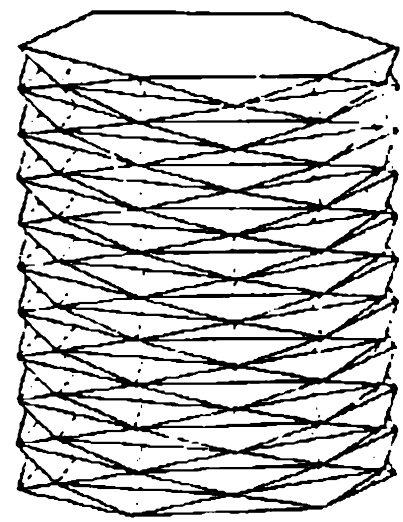


Fig. 3.21

always tends to the quantity $2\pi RH$ which is the lateral surface area of the cylinder. Depending on the way n and m are varied the total area can increase unlimitedly, tend to a finite limit different from $2\pi RH$ or have no limit at all.

In fact, simple calculations show that the area of one triangle (for given m and n) is equal to

$$R \sin \frac{\pi}{n} \sqrt{\left(\frac{H}{m}\right)^2 + R^2 \left(1 - \cos \frac{\pi}{n}\right)^2}$$

The total number of these triangles is obviously equal to $2nm$ and therefore the sum of their areas is

$$\sigma_{n,m} = 2Rn \sin \frac{\pi}{n} \sqrt{H^2 + R^2 m^2 \left(1 - \cos \frac{\pi}{n}\right)^2} \qquad (3.40)$$

If now n and m tend to infinity so that m increases faster than n^2 expression (3.40) increases unlimitedly. If n and m vary in such a way that the ratio $\frac{m}{n^2}$ tends to a finite limit q we have

$$\lim_{n,\, m \to \infty} m \left(1 - \cos \frac{\pi}{n}\right) = \lim_{n,\, m \to \infty} m 2 \sin^2 \frac{\pi}{2n} = \frac{\pi^2}{2} q$$

and, consequently,

$$\lim_{n,\, m \to \infty} \sigma_{n,m} = 2\pi R \sqrt{H^2 + \frac{\pi^4 R^2}{4} q^2}$$

Selecting q in an appropriate manner we can make the limit be equal to any number greater than $2\pi RH$ (or equal to it), i.e. to any number greater than the "true" area of the lateral surface of the cylinder. The true value of the area is obtained only in the case $q = 0$, i.e. if m increases slower than n^2.

Thus, the attempt to determine the area of a curvilinear surface by means of inscribed polyhedrons has failed even in the case of an ordinary circular cylinder. Hence, this method of defining the arc length of a curve is inapplicable to the area of a surface. This admits of a simple explanation. When the fineness of a partition of a curve (which is supposed to be smooth) is small enough the direction of the chord joining two neighbouring points of division is close to the direction of the tangent drawn at any point of the corresponding arc. But this is not the case for a surface. Indeed, a polyhedral plane area of arbitrarily small linear sizes can have all vertices lying on a smooth surface and at the same time the angle between the normal to the polyhedron and that of the surface can be large. It is apparent that such a plane element cannot serve as a good approximation to the corresponding curvilinear surface element. This is just the case in Schwarz's example: if $q \approx \frac{m}{n^2}$ is large the triangles forming the inscribed polyhedral surface are almost perpendicular to the lateral surface of the cylinder. The polyhedron composed of them forms a crinkled surface. This is why the area of such a polyhedron can be many times that of the lateral surface of the cylinder.

5. Computing Area of a Smooth Surface. In the foregoing section we have introduced the definition of the area of a curvilinear surface. We are now going to prove the existence of area for a smooth surface and deduce a formula for practical computation of the area.

Theorem 3.1. *Let a parametric equation*

$$\mathbf{r} = \mathbf{r}(u, v)$$

determine a smooth surface Σ bounded by a piecewise smooth contour. Then the surface is squarable and its area is equal to

$$\sigma = \iint\limits_D \sqrt{g_{11}g_{22} - g_{12}^2}\, du\, dv \tag{3.41}$$

where g_{11}, g_{12} and g_{22} are the ***fundamental coefficients (quantities) of the first order*** *of the surface, i.e. the coefficients of its first fundamental quadratic form, and D is the range of the variables u and v.*

Proof. Break up the surface Σ into parts Σ_i ($i = 1, 2, \ldots, n$). Choose a point M_i in each part and draw the tangent plane at it.

Next, introduce, in each tangent plane, a local coordinate system with origin at the point M_i, the normal to the surface at the point being taken as a z-axis and the tangent plane as an x, y-plane. The coordinates x, y and z of an arbitrary point of the surface Σ_i can be written as functions of u and v:

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v)^*$$

The projection of the surface Σ_i on the tangent plane at the point M_i is determined by the equations

$$x = x(u, v), \quad y = y(u, v), \quad z = 0$$

Taking advantage of the expression for the area of a plane figure in curvilinear coordinates (see Chapter 1, § 6) we can write down the area of the projection in the form

$$\pm \iint\limits_{D_i} \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} du\, dv$$

where D_i is the range of the variables u, v as the point (x, y, z) runs over the element Σ_i, the sign $+$ or $-$ being so chosen that the whole expression is positive.

The quantity

$$\pm \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}$$

can be expressed in a form irrelevant to the particular choice of the coordinate system, namely

$$\pm \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} = |[\mathbf{r}_u, \mathbf{r}_v]|^{**}$$

If the surface elements Σ_i (and, consequently, the domains D_i) are sufficiently small we have

$$\iint\limits_{D_i} |[\mathbf{r}_u, \mathbf{r}_v]|\, du\, dv = \{|[\mathbf{r}_u, \mathbf{r}_v]|\,|_{u=u_i,\, v=v_i} + \varepsilon_i\}\, d_i$$

---

* More correctly, we should have written $x = x_i(u, v)$, $y = y_i(u, v)$, $z = z_i(u, v)$ because these equations are associated with the ith coordinate system corresponding to the tangent plane and the normal at the point M_i.

** Let $\mathbf{r}_i$ be the radius vector of the point M_i in the original coordinate system in which the surface Σ has the parametric equation $\mathbf{r} = \mathbf{r}(u, v)$. Denoting the radius vector of a point M (belonging to an element Σ_i) in the local coordinate system as ρ we can write $\mathbf{r} = \mathbf{r}_i + \rho$. The vector $\mathbf{r}_i$ being considered fixed (and thus independent of u and v), we have $\mathbf{r}_u = \rho_u$, $\mathbf{r}_v = \rho_v$ and hence indeed

$$\pm \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} = |[\rho_u, \rho_v]| = |[\mathbf{r}_u, \mathbf{r}_v]|$$

where x, y (and z) are the coordinates of ρ relative to the local coordinate system.—*Tr.*

where d_i is the area of the domain D_i, u_i and v_i are the coordinates of the point M_i and $\max \varepsilon_i \to 0$ when the partition of the surface Σ is infinitely refined. Therefore, the sum of the areas of the projections of all the subdivisions Σ_i of the surface Σ on the corresponding tangent planes is equal to the expression

$$\sum_{i=1}^{n} \{|[\mathbf{r}_u, \mathbf{r}_v]|\,|_{u=u_i,\, v=v_i}\} d_i + \sum_{i=1}^{n} \varepsilon_i d_i \tag{3.42}$$

It is the limit of this expression, as the fineness of the partition of the surface tends to zero, that has been called the area of the surface. The limit exists and equals the integral

$$\iint_D |[\mathbf{r}_u, \mathbf{r}_v]|\, du\, dv$$

because the first term in (3.42) is an integral sum for the integral and the limit of the second term is zero. To complete the proof we must show that

$$|[\mathbf{r}_u, \mathbf{r}_v]| = \sqrt{g_{11}g_{22} - g_{12}^2} \tag{3.43}$$

Let ω be the angle between the vectors $\mathbf{r}_u$ and $\mathbf{r}_v$. Then

$$|[\mathbf{r}_u, \mathbf{r}_v]| = |\mathbf{r}_u|\,|\mathbf{r}_v| \sin\omega = |\mathbf{r}_u|\,|\mathbf{r}_v| \sqrt{1-\cos^2\omega} =$$
$$= \sqrt{\mathbf{r}_u^2\mathbf{r}_v^2 - \mathbf{r}_u^2\mathbf{r}_v^2\cos^2\omega} = \sqrt{g_{11}g_{22} - g_{12}^2}$$

and thus the theorem has been proved.

Note 1. We have already dealt with the vector $[\mathbf{r}_u, \mathbf{r}_v]$ (see § 3, Sec. 5). As has been shown, its projections are

$$A = \begin{vmatrix} \dfrac{\partial y}{\partial u} & \dfrac{\partial z}{\partial u} \\ \dfrac{\partial y}{\partial v} & \dfrac{\partial z}{\partial v} \end{vmatrix}, \quad B = \begin{vmatrix} \dfrac{\partial z}{\partial u} & \dfrac{\partial x}{\partial u} \\ \dfrac{\partial z}{\partial v} & \dfrac{\partial x}{\partial v} \end{vmatrix} \quad \text{and} \quad C = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

(where x, y and z are now the coordinates of $\mathbf{r}$ in the original coordinate system) and hence the length of the vector is equal to

$$\sqrt{A^2 + B^2 + C^2}$$

Consequently, formula (3.41) for the area of a surface can be rewritten as follows:

$$\sigma = \iint_D \sqrt{A^2 + B^2 + C^2}\, du\, dv \tag{3.41$'$}$$

Note 2. The geometric meaning of formula (3.41) lies in the fact that the element of integration $\sqrt{g_{11}g_{22} - g_{12}^2}\, du\, dv$ coincides, to within infinitesimals of higher order, with the area of an "infinitesimal parallelogram" cut out of the surface Σ by two pairs of coordi-

nate curves $u = u_0$, $u = u_0 + du$ and $v = v_0$, $v = v_0 + dv$ drawn infinitely close to each other (see Fig. 3.22). In fact, the vertices P_0, P_1 and P_2 of the parallelogram have the curvilinear coordinates

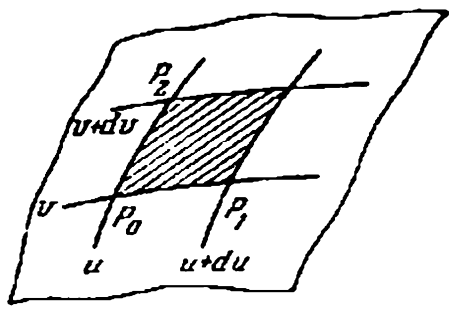


Fig. 3.22

(u_0, v_0), $(u_0 + du, v_0)$ and $(u_0, v_0 + dv)$, respectively. Therefore, we have, to within infinitesimals of order higher than the first, the relations

$$\overline{P_0P_1} = \mathbf{r}_u du \quad \text{and} \quad \overline{P_0P_2} = \mathbf{r}_v dv$$

The area $d\sigma$ of the parallelogram constructed on the vectors $\overline{P_0P_1}$ and $\overline{P_0P_2}$ is equal to the absolute value of their vector product:

$$d\sigma = |\,[\mathbf{r}_u, \mathbf{r}_v]\,|\, du\, dv$$

Finally, by virtue of formula (3.43), the last expression can be put down as

$$d\sigma = \sqrt{g_{11}g_{22} - g_{12}^2}\, du\, dv$$

Let us consider some important special cases of formula (3.41). If a surface Σ is given by an equation

$$z = f(x, y)$$

expressing z as an explicit function of x and y we can write, as has been shown (see Sec. 2, example 4), the formulas

$$g_{11} = 1 + f_x'^2, \quad g_{12} = f_x' f_y' \quad \text{and} \quad g_{22} = 1 + f_y'^2$$

whence

$$\sqrt{g_{11}g_{22} - g_{12}^2} = \sqrt{1 + f_x'^2 + f_y'^2}$$

Thus, the area of a surface $z = f(x, y)$ is expressed by the formula

$$\sigma = \iint\limits_D \sqrt{1 + f_x'^2 + f_y'^2}\, dx\, dy \tag{3.44}$$

where D is (for this particular case) the projection of the entire surface Σ on the x, y-plane.

Note 1. Since we have

$$\sqrt{1 + f_x'^2 + f_y'^2} = \frac{1}{\cos(\mathbf{N}, z)}$$

(see § 3, Sec. 5), formula (3.44) can also be written in the form

$$\sigma = \iint\limits_D \frac{dx\,dy}{\cos(N, z)}$$

This formula admits of a simple geometric interpretation: the area of a surface element is equal to the area of its projection on the x, y-plane divided by the cosine of the angle between the normal to the element and that to the x, y-plane.

Note 2. If a surface Σ is composed of a finite number of pieces each of which is representable by an equation of the form $z = f(x, y)$ its area can be found by applying formula (3.44) separately to each of the pieces.

Example. Find the area of the part of the paraboloid $z = x^2 + y^2$ cut out by the cylinder $x^2 + y^2 = a^2$.

Solution. The sought-for area is equal to

$$\sigma = \iint\limits_{x^2+y^2\leqslant a^2} \sqrt{1+4x^2+4y^2}\,dx\,dy$$

Passing to polar coordinates we obtain

$$\sigma = \int_0^{2\pi} d\varphi \int_0^a \sqrt{4r^2+1}\,r\,dr = \frac{1}{12}\int_0^{2\pi} [(4r^2+1)^{3/2}]\Big|_0^a d\varphi =$$

$$= \frac{1}{12}\int_0^{2\pi} [(4a^2+1)^{3/2}-1]\,d\varphi = \frac{\pi}{6}[(4a^2+1)^{3/2}-1]$$

Suppose now that a surface is determined by an equation $F(x, y, z) = 0$ expressing an implicit functional relationship between z and the variables x and y. If the surface is such that it is possible to solve the equation in z, which is equivalent to the requirement that every vertical (i.e. parallel to the z-axis) straight line has at most one point in common with the surface, we can apply the rules for differentiation of an implicit function and thus write

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}}, \qquad \frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}$$

Substituting these expressions for f'_x and f'_y into formula (3.44) we derive

$$\sigma = \iint\limits_D \frac{\sqrt{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 + \left(\frac{\partial F}{\partial z}\right)^2}}{\left|\frac{\partial F}{\partial z}\right|}\,dx\,dy \qquad (3.45)$$

Here again, as in formula (3.44), the integrand is nothing but the reciprocal of the cosine of the angle between the normal to the surface and the z-axis.

Exercise. Determine the area of the part of the conical surface $x^2 + y^2 - z^2 = 0$ lying inside the cylinder $x^2 + y^2 = a^2$.

Answer. $\sigma = 2\sqrt{2}\,\pi a^2$.

§ 5. CURVATURE OF CURVES ON A SURFACE. SECOND FUNDAMENTAL QUADRATIC FORM OF A SURFACE

In the foregoing sections we deduced the formulas for computing lengths of curves on a surface, angles between the curves and areas of surfaces. But these quantities do not completely characterize the shape of the surface. For instance, a cylinder and a plane are obviously different surfaces although a cylinder can be rolled out on a plane so that all the angles, lengths and areas are preserved. To investigate the shape of a surface we shall apply the following method: we draw all the possible planes passing through the normal to the surface at a given point and consider the shape of the sections, i.e. the plane curves (called **normal sections**) thus obtained.

1. Normal Sections of a Surface and Their Curvature. Let us take a surface Σ determined by an equation

$$\mathbf{r} = \mathbf{r}(u, v)$$

Here and henceforward the vector function $\mathbf{r}(u, v)$ will be supposed to be doubly continuously differentiable. Choose a point M_0

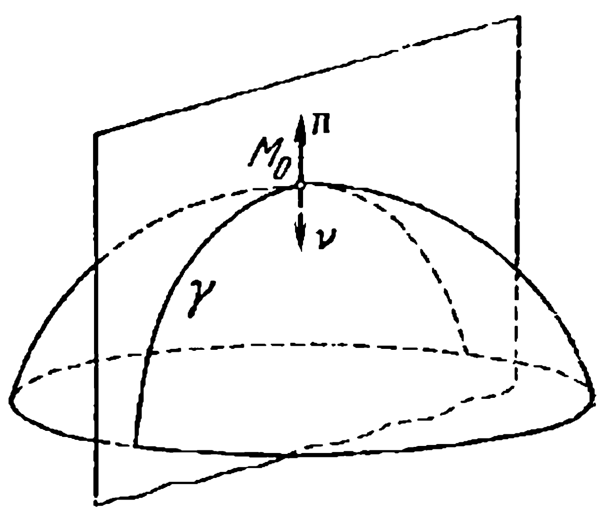


Fig. 3.23

on the surface and define a certain direction on the normal to the surface Σ at the point M_0, i.e. draw a unit vector $\mathbf{n}$ along the normal line. Let γ be one of the normal sections passing through the point M_0. Then the curve γ lies in a plane passing through unit normal vector $\mathbf{n}$ to Σ at the point M_0 (Fig. 3.23). Thus, γ is a plane curve,

its shape in the vicinity of the point M_0 being completely characterized by its curvature k at the point and by its direction of concavity (relative to the chosen direction of the normal, at the point M_0, specified by the vector $\mathbf{n}$). To compute the curvature of the curve γ write the equation of the curve in the form

$$\mathbf{r}=\mathbf{r}(u(l), v(l)) \tag{3.46}$$

where l is its arc length and apply the first Frenet-Serret formula

$$\frac{d\boldsymbol{\tau}}{dl}=\frac{d^2\mathbf{r}}{dl^2}=k\boldsymbol{\nu}$$

It follows that

$$k=\left(\frac{d^2\mathbf{r}}{dl^2}, \boldsymbol{\nu}\right) \tag{3.47}$$

The unit vector $\boldsymbol{\nu}$ is apparently in the direction of the normal to the surface Σ at the point M_0 and, consequently, it coincides with

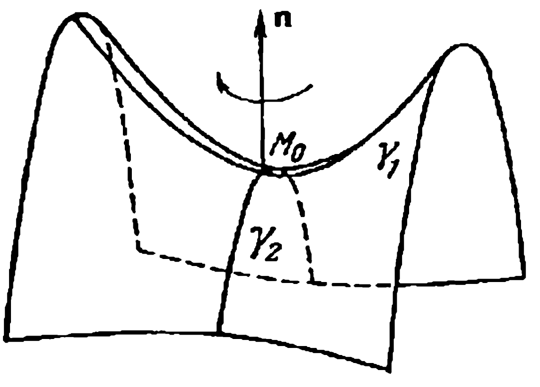


Fig. 3.24

$\mathbf{n}$ in case the direction of concavity of the section γ coincides with the direction chosen along the normal to Σ or differs from $\mathbf{n}$ in its sign if these directions are opposite. In order to take into account both the value of the curvature and the direction of concavity of the section γ we introduce the quantity

$$\widetilde{k}=\left(\frac{d^2\mathbf{r}}{dl^2}, \mathbf{n}\right) \tag{3.48}$$

which will be referred to as the **normal curvature of the surface Σ at the point M_0 in the direction of the section γ**. It is clear that $k=|\widetilde{k}|$. If the plane in which the section γ lies is rotated about the vector $\mathbf{n}$ the normal curvature $\widetilde{k}=\widetilde{k}(\gamma)$ may vary. Its variation indicates not only the shape of the normal section but also the direction of its concavity. For instance, if a surface is of the form of a saddle in the vicinity of the point M_0, as shown in Fig. 3.24, we have a positive normal curvature $\widetilde{k_1}$ for the section γ_1 since the vector $\boldsymbol{\nu}_1$ of the principal normal to γ_1 coincides with $\mathbf{n}$ and a negative normal curvature $\widetilde{k}_2$ for the section γ_2 since $\boldsymbol{\nu}_2=-\mathbf{n}$.

In what follows we shall always consider normal curvature (3.48) but not the curvature defined by formula (3.47). This normal curvature will be denoted by the letter k instead of $\tilde{k}$.

The quantity

$$R = \frac{1}{k}$$

is known as the **radius of normal curvature** of the surface Σ (at the corresponding point and in the given direction). The nonnegative quantity $|R|$ is obviously the radius of curvature of the corresponding normal section. Since k may vanish the quantity R may assume infinite values.

Let us now derive a formula for computing the normal curvature k. For this purpose we take advantage of equation (3.46) of the curve γ and calculate $\frac{d^2\mathbf{r}}{dl^2}$. For brevity, we introduce the notation

$$\mathbf{r}_{uu} = \frac{\partial^2 \mathbf{r}}{\partial u^2}, \qquad \mathbf{r}_{uv} = \frac{\partial^2 \mathbf{r}}{\partial u\, \partial v}, \qquad \mathbf{r}_{vv} = \frac{\partial^2 \mathbf{r}}{\partial v^2}$$

From equation (3.46) of the curve γ we deduce

$$\frac{d^2\mathbf{r}}{dl^2} = \frac{d}{dl}\left(\mathbf{r}_u \frac{du}{dl} + \mathbf{r}_v \frac{dv}{dl}\right) =$$
$$= \mathbf{r}_{uu}\left(\frac{du}{dl}\right)^2 + 2\mathbf{r}_{uv}\frac{du}{dl}\frac{dv}{dl} + \mathbf{r}_{vv}\left(\frac{dv}{dl}\right)^2 + \mathbf{r}_u \frac{d^2u}{dl^2} + \mathbf{r}_v \frac{d^2v}{dl^2} \qquad (3.49)$$

The vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ lie in the tangent plane and hence they are orthogonal to $\mathbf{n}$, that is

$$(\mathbf{r}_u, \mathbf{n}) = (\mathbf{r}_v, \mathbf{n}) = 0$$

Therefore, substituting expression (3.49) for $\frac{d^2\mathbf{r}}{dl^2}$ into formula (3.48) we obtain

$$k = \frac{1}{R} = \left(\frac{d^2\mathbf{r}}{dl^2}, \mathbf{n}\right) =$$
$$= (\mathbf{r}_{uu}, \mathbf{n})\left(\frac{du}{dl}\right)^2 + 2(\mathbf{r}_{uv}, \mathbf{n})\frac{du}{dl}\frac{dv}{dl} + (\mathbf{r}_{vv}, \mathbf{n})\left(\frac{dv}{dl}\right)^2 \qquad (3.50)$$

2. Second Fundamental Quadratic Form of a Surface. Let us transform formula (3.50) for the normal curvature to another form which is more convenient. Introduce the notation

$$b_{11} = (\mathbf{r}_{uu}, \mathbf{n}), \quad b_{12} = (\mathbf{r}_{uv}, \mathbf{n}), \quad b_{22} = (\mathbf{r}_{vv}, \mathbf{n}) \qquad (3.51)$$

and rewrite equality (3.50) as follows:

$$k = \frac{1}{R} = \frac{b_{11}\, du^2 + 2b_{12}\, du\, dv + b_{22}\, dv^2}{dl^2} \qquad (3.52)$$

Here we have the expression dl^2, i.e. the first fundamental quadratic form of the surface, in the denominator. The numerator is also a quadratic form (in the variables du and dv). It is called the **second fundamental quadratic form of the surface** and plays a very important role (together with the first fundamental quadratic form) in the theory of surfaces. In what follows, the second fundamental quadratic form of a surface will be denoted by the symbol φ_2.
Thus, we have

$$\varphi_2 = b_{11}\,du^2 + 2b_{12}\,du\,dv + b_{22}\,dv^2$$

where b_{11}, b_{12} and b_{22} are determined by relations (3.51).

Example. Take a surface defined by an equation

$$z = f(x, y)$$

or, in vector form,

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + f(x, y)\,\mathbf{k}$$

Here we have

$$\mathbf{r}_{xx} = f''_{xx}\mathbf{k}, \qquad \mathbf{r}_{xy} = f''_{xy}\mathbf{k} \quad \text{and} \quad \mathbf{r}_{yy} = f''_{yy}\mathbf{k}$$

Consequently,

$$b_{11} = f''_{xx}\cos(\mathbf{n}, z), \qquad b_{12} = f''_{xy}\cos(\mathbf{n}, z) \quad \text{and} \quad b_{22} = f''_{yy}\cos(\mathbf{n}, z)$$

that is

$$\varphi_2 = (f''_{xx}\,dx^2 + 2f''_{xy}\,dx\,dy + f''_{yy}\,dy^2)\cos(\mathbf{n}, z) \tag{3.53}$$

Thus, in this case the second fundamental quadratic form is, to within the factor $\cos(\mathbf{n}, z)$, the sum of the second-order terms in the expansion of the function $z = f(x, y)$ by Taylor's formula.

Note. As has been shown, the first fundamental quadratic form of a surface determines its "metric", i.e. such quantities as lengths, angles and areas which are found by means of the form. The computation of these quantities is in fact based on the replacement, in the first approximation, of an infinitesimal surface element by the corresponding element of its tangent plane. The second fundamental quadratic form of a surface characterizes the measure of the rate at which the surface turns out of the tangent plane drawn through its point in the vicinity of the point.

To prove this, let us find the distance from a point M, of a surface Σ, lying close to a given point M_0, through which the tangent plane to Σ is drawn, to the plane (see Fig. 3.25). Consider a normal section passing through the points M_0 and M. The sought-for distance is apparently equal to the distance MP from M to the tangent line to the curve γ. This distance is equal, to within infinitesimals of higher order (see § 2, Sec. 6), to

$$\frac{1}{2}k\,dl^2 = \frac{1}{2}(b_{11}\,du^2 + 2b_{12}\,du\,dv + b_{22}\,dv^2)$$

the sign of the last expression indicating the direction in which the surface is turning out of the tangent plane.

It is possible to give a definition of the second fundamental quadratic form of a surface (equivalent to the definition given above)

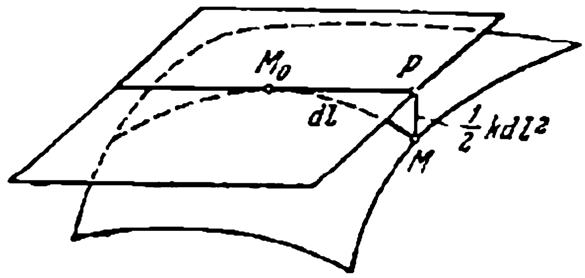


Fig. 3.25

proceeding from the problem of calculating the distance from a point of the surface to the tangent plane drawn through another point taken close to the former.

Exercises

1. Prove that the second fundamental quadratic form of a plane is identically equal to zero for all the possible parametric representations of the plane.

2. Find the second fundamental quadratic form of a torus in coordinates φ and ψ (see Example 1 in § 3, Sec. 1).

3. Dupin Indicatrix. The radius of normal curvature $R = \frac{1}{k}$ corresponding to a normal section γ at a point M_0 depends on the direction in which the section γ is drawn. To represent the dependence

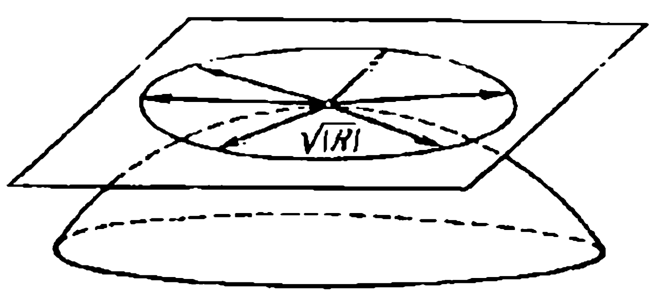


Fig. 3.26

in a visual manner we can apply the following technique. Lay off, from the point M_0, in all the directions on the tangent plane, a radius vector ρ whose length is equal to $\sqrt{|R|}$ where R is the radius of normal curvature of the surface in the corresponding direction. The vector can obviously be written in the form

$$\rho = \sqrt{|R|}\,\tau$$

where τ is the unit tangent vector to the normal section in question.

The locus of the tips of the vectors is a curve lying in the tangent plane to the surface Σ at the point M_0 (Fig. 3.26). It is called Du-

pin's* **indicatrix** of the surface Σ at the point. Let us deduce the equation of the Dupin indicatrix.

Take the vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ as base vectors of the coordinate system in the tangent plane. We have

$$\tau = \frac{d\mathbf{r}}{dl} = \mathbf{r}_u \frac{du}{dl} + \mathbf{r}_v \frac{dv}{dl}$$

and therefore

$$\rho = \sqrt{|R|} \frac{du}{dl} \mathbf{r}_u + \sqrt{|R|} \frac{dv}{dl} \mathbf{r}_v$$

that is each point of the Dupin indicatrix has the coordinates

$$\xi = \sqrt{|R|} \frac{du}{dl} \quad \text{and} \quad \eta = \sqrt{|R|} \frac{dv}{dl}$$

relative to the basis we have chosen.

Next we take advantage of the relation

$$\frac{1}{R} = b_{11} \left(\frac{du}{dl}\right)^2 + 2b_{12} \frac{du}{dl} \frac{dv}{dl} + b_{22} \left(\frac{dv}{dl}\right)^2$$

Multiplying it by $|R|$ we see that

$$b_{11} \left(\sqrt{|R|} \frac{du}{dl}\right)^2 + 2b_{12} \left(\sqrt{|R|} \frac{du}{dl}\right) \left(\sqrt{|R|} \frac{dv}{dl}\right) +$$
$$+ b_{22} \left(\sqrt{|R|} \frac{dv}{dl}\right)^2 = \pm 1$$

which implies that ξ and η satisfy the equation

$$b_{11}\xi^2 + 2b_{12}\xi\eta + b_{22}\eta^2 = \pm 1 \tag{3.54}$$

This is an equation of a central curve of the second order with centre at the origin of coordinates.**

Thus, Dupin's indicatrix is a central second-degree curve with centre at the corresponding point of the surface.***

4. Principal Directions and Principal Curvatures of a Surface. Equation of Euler. The Dupin indicatrix being a central curve of the second order, we can pass to its principal axes, i.e. replace the base vectors $\mathbf{r}_u$ and $\mathbf{r}_v$, if necessary, by a pair of unit base vectors lying in the tangent plane which are mutually orthogonal and such that the equation of the Dupin indicatrix in the new coordinates does not contain the terms with the product of the coordinates. The new base vectors must be in the directions of the principal axes of the Dupin indicatrix. We shall call the latter the **principal directions** of the surface (at the point in question).

---

* Dupin, François Pierre Charles (1784-1873), a French mathematician.

** This is implied by the fact that there are no first-order terms in the equation.

*** More precisely, there are two such curves here, namely $b_{11}\xi^2 + 2b_{12}\xi\eta + b_{22}\eta^2 = 1$ and $b_{11}\xi^2 + 2b_{12}\xi\eta + b_{22}\eta^2 = -1$ whose equations only differ in the signs of their constant terms. For more detail concerning the shape of the Dupin indicatrix see Sec. 7.

For such a choice of the coordinate system in the tangent plane, the equation of the Dupin indicatrix takes the form

$$px^2 + qy^2 = \pm 1 \tag{3.55}$$

Let φ be the angle between the principal direction taken as the direction of the x-axis and an arbitrary normal section. Then we obviously have

$$x = \sqrt{|R|}\cos\varphi, \quad y = \sqrt{|R|}\sin\varphi$$

where R is the radius of curvature of the corresponding normal section. Substituting these expressions of x and y into equation (3.55) and bearing in mind that the right-hand side of the equation is equal to the ratio of $|R|$ to R we obtain

$$p\cos^2\varphi + q\sin^2\varphi = \frac{1}{R} = k \tag{3.56}$$

Denote by $k_1 = \frac{1}{R_1}$ and $k_2 = \frac{1}{R_2}$ the normal curvatures corresponding to the principal directions of the Dupin indicatrix at the point under consideration. These quantities are referred to as the **principal curvatures** of the surface at the point. The principal directions are determined by the values $\varphi = 0$ and $\varphi = \frac{\pi}{2}$ in the coordinate system we have chosen in the tangent plane. Consequently, we have

$$k_1 = p, \quad k_2 = q$$

Therefore equality (3.60) takes the form

$$k = k_1\cos^2\varphi + k_2\sin^2\varphi \tag{3.57}$$

or

$$\frac{1}{R} = \frac{1}{R_1}\cos^2\varphi + \frac{1}{R_2}\sin^2\varphi \tag{3.57'}$$

Formula (3.57) or (3.57') is known as the **equation of Euler***. It expresses the normal curvature in an arbitrary direction in terms of the principal curvatures. From the equation of Euler it immediately follows that the principal curvatures are the extremal values of the normal curvature. Indeed, if $k_1 = k_2$, the quantity k is independent of φ, and all the directions can be regarded as being extremal in this case.** But if $k_1 \neq k_2$ we can put, for definiteness, $k_1 > k_2$, and then $k_1 - k_2 > 0$ and the equation of Euler can be rewritten as

$$k = (k_1 - k_2)\cos^2\varphi + k_2(\cos^2\varphi + \sin^2\varphi) =$$
$$= (k_1 - k_2)\cos^2\varphi + k_2$$

which shows that $k_1 \geqslant k \geqslant k_2$ for every φ.

* Euler, Leonard (1707-1783), a great Russian mathematician (a Swiss by birth).

** A point on a surface at which $k_1 = k_2 \neq 0$ is said to be an umbilical (circular) point of the surface. It can be shown that the only surface whose all points are umbilical is the sphere.

The extremal properties of the principal curvatures provide a convenient practical method for computing them.

5. Determining Principal Curvatures. Equation of Euler (3.57) makes it possible to visualize the dependence of the normal curvature $k(\varphi)$ on the direction specified by the angle φ. This functional relationship is represented graphically in Fig. 3.27. It shows that

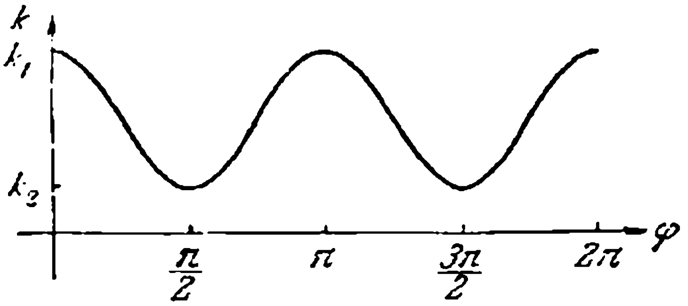


Fig. 3.27

for every given k_0, $k_1 > k_0 > k_2$, there exist four values of the angle φ for which $k(\varphi) = k_0$. Since the angles which differ by π define the same direction, there are two normal sections corresponding to each k_0 for which the normal curvature is equal to k_0. But when $k_0 = k_1$ or $k_0 = k_2$ the two normal sections merge into one.

In other words, the principal curvatures are the values of the normal curvature such that to each of them there corresponds one and only one normal section of the surface. Formula (3.52) defining the normal curvature as a function of the direction can be rewritten as follows:

$$(b_{11} - kg_{11})\, du^2 + 2\,(b_{12} - kg_{12})\, du\, dv + (b_{22} - kg_{22})\, dv^2 = 0$$

Now, dividing by dv^2 and putting $\frac{du}{dv} = t$ (where the value t determines the direction of the section) we obtain

$$(b_{11} - kg_{11})\, t^2 + 2\,(b_{12} - kg_{12})\, t + (b_{22} - kg_{22}) = 0 \qquad (3.58)$$

According to the above discussion, this quadratic equation (in t) has a single root and not two distinct roots if and only if the parameter k entering into (3.58) takes on the values of the principal curvatures. Then the corresponding values of the single root t determine the principal directions. Furthermore, for this being so it is necessary and sufficient that the discriminant of equation (3.58) turn into zero.

Thus, to find the principal curvatures we must solve, in k, the following equation:

$$(b_{12} - kg_{12})^2 - (b_{11} - kg_{11})\,(b_{22} - kg_{22}) = 0 \qquad (3.59)$$

or

$$\begin{vmatrix} b_{11} - kg_{11} & b_{12} - kg_{12} \\ b_{12} - kg_{12} & b_{22} - kg_{22} \end{vmatrix} = 0 \qquad (3.59')$$

6. Total Curvature and Mean Curvature. In many cases it is more convenient to consider the product

$$K = k_1 k_2 \tag{3.60}$$

and the half-sum

$$H = \frac{1}{2}(k_1 + k_2) \tag{3.61}$$

instead of the principal curvatures k_1 and k_2. The quantity K is called the **total** (or **Gaussian**) **curvature** of the surface, and H is referred to as its **mean curvature**.

Quadratic equation (3.59′) immediately implies the formulas

$$K = \frac{b_{11}b_{22} - b_{12}^2}{g_{11}g_{22} - g_{12}^2}, \qquad H = \frac{g_{11}b_{22} - 2g_{12}b_{12} + g_{22}b_{11}}{2(g_{11}g_{22} - g_{12}^2)} \tag{3.62}$$

Example. Compute the total and the mean curvatures for the hyperbolic paraboloid $z = x^2 - y^2$.

Solution. We have $g_{11} = 1 + 4x^2$, $g_{12} = -4xy$, $g_{22} = 1 + 4y^2$, $b_{11} = 2$, $b_{12} = 0$ and $b_{22} = -2$. Hence,

$$K = -\frac{4}{1 + 4x^2 + 4y^2}, \qquad H = \frac{4(x^2 - y^2)}{1 + 4x^2 + 4y^2}$$

In particular, we have $K = -4$ and $H = 0$ at the origin.

7. Classification of Points on a Surface. We have attributed a certain plane curve, namely the Dupin indicatrix, to each point M_0 of a surface Σ defined by equations involving doubly differentiable functions. As has been shown, the equation of Dupin's indicatrix can be transformed to the form

$$k_1 x^2 + k_2 y^2 = \pm 1 \tag{3.63}$$

where k_1 and k_2 are the principal curvatures of the surface at the point M_0. The type of curve (3.63) depends on the sign of the product $k_1 k_2$. Let us consider the possible cases.

(1) $k_1 k_2 > 0$. We can put, without loss of generality, $k_1 > 0$ and $k_2 > 0$, because, if otherwise, we can reverse the direction of the normal vector **n** and thus change the signs of k_1 and k_2 to the opposite. Equation (3.63) determines an ellipse for $k_1 > 0$ and $k_2 > 0$ if we have $+1$ on its right-hand side and does not define any curve at all when we have -1.

The points for which $k_1 k_2 > 0$ (i.e. when the Dupin indicatrix is an ellipse) are called **elliptic points** of the surface.

(2) $k_1 k_2 < 0$. In this case equation (3.63) determines a hyperbola or, more precisely, two hyperbolas with common asymptotes. One of them corresponds to the term $+1$ on the right-hand side and the other to -1. A point at which $k_1 k_2 < 0$ (when the Dupin indicatrix is a pair of hyperbolas) is called a **hyperbolic point**.

(3) $k_1k_2 = 0$. If one of the principal curvatures is different from zero equation (3.63) determines a pair of straight lines intersecting at the point. The points at which $k_1k_2 = 0$ (but one of the principal curvatures is nonzero) are said to be **parabolic**.

If $k_1 = k_2 = 0$ the notion of Dupin's indicatrix becomes senseless. A point for which $k_1 = k_2 = 0$ is referred to as a planar point of the surface.

Thus, the type of the point is specified by the total curvature $K = k_1k_2$ at the point. Since

$$K = \frac{b_{11}b_{22} - b_{12}^2}{g_{11}g_{22} - g_{12}^2}$$

and the quantity $g_{11}g_{22} - g_{12}^2$ is always positive, the type of the point is determined by the sign of the discriminant of the second fundamental quadratic form.

We can easily visualize the shape of a surface in the vicinity of its point belonging to each type. Let M_0 be an elliptic point. Then

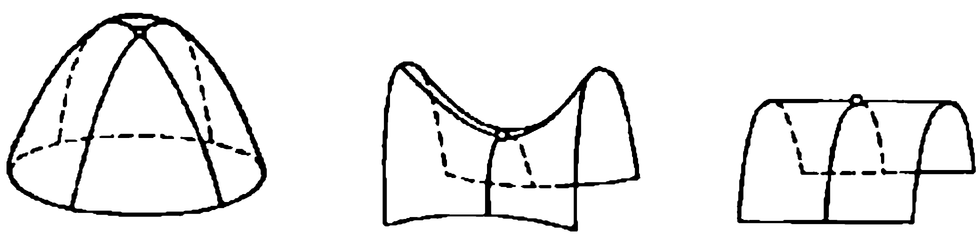


Fig. 3.28 (a) (b) (c)

k_1 and k_2 are of the same sign and hence, by virtue of the equation of Euler, all the normal curvatures have the same sign at the point. This means, geometrically, that all the normal sections at the point have the same direction of concavity. In the vicinity of an elliptic point the surface resembles a piece of an ellipsoid and looks as is shown in Fig. 3.28a.

Consider now a hyperbolic point. The principal curvatures are of opposite signs at the point. Therefore in this case there are normal sections of different directions of concavity. The surface is of the shape of a saddle in the vicinity of such a point (see Fig. 3.28b).

The structure of a surface in the vicinity of its parabolic point can be of a more complicated nature. In this case there is a direction in which the normal curvature is equal to zero, and the normal curvatures are nonzero and of the same sign in all the other directions. A typical example of a parabolic point is any point of an ordinary circular cylinder (see Fig. 3.28c) but there are many other possible configurations which we shall not discuss here.

Consider an example. Let a surface be determined by an equation

$$z = f(x, y)$$

and let the well known necessary conditions for extremum be fulfilled at a point (x_0, y_0), i.e. $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$. Then the normal

to the surface at the point coincides, in its direction, with the z-axis, and, as it can be easily shown by means of simple calculations, the coefficients of the second fundamental quadratic form at the point are

$$b_{11}=f''_{xx},\quad b_{12}=b_{21}=f''_{xy},\quad b_{22}=f''_{yy}$$

Consequently, we have

$$b_{11}b_{22}-b_{12}^2=f''_{xx}f''_{yy}-f''^2_{xy} \tag{3.64}$$

We see that the type of the point is determined by the sign of expression (3.64). But, as is well known, the sign of the expression specifies the existence or nonexistence of an extremum at the point. Thus, we have established the following relationship between the type of the point and the existence or nonexistence of an extremum at it:

elliptic point. The condition $f''_{xx}f''_{yy}-f''^2_{xy}>0$ holds and there is an extremum;

hyperbolic point. The condition $f''_{xx}f''_{yy}-f''^2_{xy}<0$ is then fulfilled and there is no extremum;

parabolic point. The condition $f''_{xx}f''_{yy}-f''^2_{xy}=0$ which takes place here indicates the case when the question of an extremum at the point (x_0, y_0) remains open and requires further investigation.

Exercise. Determine the type of the points lying on the following surfaces: (1) an ellipsoid, (2) a hyperboloid of two sheets, (3) a hyperboloid of one sheet, (4) an elliptic paraboloid and (5) a hyperbolic cylinder.

8. The First and the Second Fundamental Quadratic Forms as Invariants of a Surface. We have introduced the first fundamental quadratic form of a surface and shown that it determines lengths, angles and areas on the surface. Furthermore, we have proved that the second fundamental quadratic form specifies the shape of the surface in the vicinity of each point. Now it is natural to ask as to what extent a surface is determined by its two fundamental quadratic forms. The answer to the question is given by the following theorem.

Theorem 3.2. *If it is possible to introduce a curvilinear coordinate system u, v on a surface Σ and a system u^*, v^* on a surface Σ^* so that at the points where $u=u^*$ and $v=v^*$ the corresponding fundamental quadratic forms also coincide (in the sense that the equalities*

$$g_{11}=g^*_{11},\quad g_{12}=g^*_{12},\quad g_{22}=g^*_{22},$$
$$b_{11}=b^*_{11},\quad b_{12}=b^*_{12},\quad b_{22}=b^*_{22}$$

hold at these points) the two surfaces are congruent, i.e. they can only differ in their position in space.

Thus, the first and the second fundamental quadratic forms of a surface play the same role for surfaces as the intrinsic equations for curves, and hence they form a complete system of invariants which uniquely specifies the surface to within its position in space.

We shall not present the proof of the theorem because it can be found in many courses in differential geometry.*

§ 6. INTRINSIC PROPERTIES OF A SURFACE

1. Applicable Surfaces. Necessary and Sufficient Condition for Applicability. In the foregoing sections we regarded a surface as a rigid body which can move in space but cannot change its shape. But it is sometimes convenient to consider a surface as an inextensible but absolutely flexible film. This leads to studying the properties of a surface which do not vary as the surface is subjected to a bending, i.e. to a deformation which is not connected with stretching or shrinking.

If a surface can be made coincident with another surface by means of a bending, the surfaces are said to be **applicable (isometric)**.

In other words, two surfaces are called applicable if it is possible to establish a one-to-one correspondence between their points so that the curves, lying on them, which are transformed into each other by the correspondence, are of the same length.

It seems natural to raise the question as to what are the necessary and sufficient conditions for two surfaces to be applicable. The answer is given by the following theorem.

Theorem 3.3. For two surfaces Σ and Σ^* to be applicable it is necessary and sufficient that it be possible to introduce a parametrization of the surfaces by means of the same parameters u and v so that their fundamental quantities of the first order (the coefficients of their first fundamental quadratic forms) should coincide at the points $M \in \Sigma$ and $M^* \in \Sigma^*$ having the same values of the coordinates u and v.

Proof. If the condition of the theorem holds we can establish a one-to-one correspondence between the points of the surfaces having the same coordinates u and v, and then their fundamental coefficients of the first order will coincide at the corresponding points:

$$g_{11} = g_{11}^*, \quad g_{12} = g_{12}^*, \quad g_{22} = g_{22}^*$$

This makes it possible to introduce parametric representations of the curves, lying on the surfaces, which are mapped onto each other, by means of a common parameter t such that the values of the parameter are the same at the points of the curves which correspond

* For instance, see [12], p. 188.

to each other. This implies

$$\int_{t_1}^{t_2} \sqrt{g_{11}\left(\frac{du}{dt}\right)^2 + 2g_{12}\frac{du}{dt}\frac{dv}{dt} + g_{22}\left(\frac{dv}{dt}\right)^2}\, dt =$$

$$= \int_{t_1}^{t_2} \sqrt{g_{11}^*\left(\frac{du}{dt}\right)^2 + 2g_{12}^*\frac{du}{dt}\frac{dv}{dt} + g_{22}^*\left(\frac{dv}{dt}\right)^2}\, dt \qquad (3.65)$$

i.e. the arcs are of equal lengths.

Conversely, if two surfaces Σ and Σ^* are applicable they can be represented parametrically with the help of the same parameters by introducing an arbitrary coordinate system u, v on the surface Σ and attributing to each point $M^* \in \Sigma^*$ the values of the coordinates u and v of the point $M \in \Sigma$ to which M^* corresponds. Take now an arbitrary curve on the surface Σ and the corresponding curve on the surface Σ^* and introduce a parameter t on them in such a way that the points which coincide when the surfaces are applied to each other have the same values of the parameter. The arc lengths of the curves being equal, we can write relation (3.65) for them. Since the relation must hold for all the possible values t_1 and t_2 of the parameter it follows that

$$g_{11}\, du^2 + 2g_{12}\, du\, dv + g_{22}\, dv^2 = g_{11}^*\, du^2 + 2g_{12}^*\, du\, dv + g_{22}^*\, dv^2$$

The last equality is an identity in du and dv since it is fulfilled for any two curves corresponding to each other and passing through any point in any direction. An identical equality of two quadratic forms implies the coincidence of their coefficients and hence

$$g_{11} = g_{11}^*, \quad g_{12} = g_{12}^*, \quad g_{22} = g_{22}^*$$

which is what we set out to prove.

2. Intrinsic Properties of a Surface. The description of the properties of a surface which are invariable under a bending (i.e. are preserved under an arbitrary **isometric, length preserving, mapping**) constitutes the **intrinsic geometry of the surface**. Such properties are referred to as **intrinsic (absolute) properties** of the surface. We have proved that two surfaces are applicable if and only if it is possible to introduce a first fundamental quadratic form common for them. Hence, a property belongs to the intrinsic geometry of a surface if and only if it is expressible in terms of its first fundamental quadratic form.* Thus, *the intrinsic geometry of a surface is determined by its first fundamental quadratic form*. Consequently, the lengths of the curves lying on a surface are relevant to its intrin-

* Here we mean, of course, the properties which are related to the surface itself but not to the particular way of introducing parameters on it.

sic geometry. Further, since the angle between two curves on a surface and the area of a surface are expressible in terms of the fundamental coefficients of the first order (see Secs. 2 and 4 in § 4), these quantities also belong to the intrinsic geometry of the surface.

A remarkable fact is that the intrinsic geometry of a surface includes its total (Gaussian) curvature K. Indeed, there is a formula obtained by Gauss which expresses the total curvature in curvilinear orthogonal coordinates in the form

$$K=-\frac{1}{2\sqrt{g_{11}g_{22}}}\left\{\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{g_{11}g_{22}}}\frac{\partial g_{22}}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{1}{\sqrt{g_{11}g_{22}}}\frac{\partial g_{11}}{\partial v}\right)\right\}$$

which involves only the fundamental quantities of the first order. At the same time, neither the mean curvature nor the principal curvatures are preserved under an arbitrary isometric transformation.

The term "intrinsic geometry" means that the properties under consideration, preserved under an arbitrary isometric mapping, pertain merely to the surface, not to its position in the surrounding space.

We can illustrate this by means of the following "mental" experiment. Imagine that there are some intelligent creatures inhabiting a two-dimensional surface and that they cannot leave it and go out into the surrounding space. They can construct the geometry of their "world", introduce the notion of a "straight line" passing through two points by defining it as the shortest curve entirely lying on the surface that joins the points (for instance, in the case of a sphere such "straight lines" are the arcs of the great circles) and so on. They can introduce the notions of a "triangle", "polygon" etc. and study the properties of these figures (without going out into the space surrounding the surface). These hypothetical creatures cannot distinguish between this surface and any other surface applicable to it.*

The geometry thus obtained is nothing but the intrinsic geometry of the surface. For instance, the intrinsic geometry of the plane is ordinary planimetry studied in elementary geometrical courses. But all the theorems of planimetry remain true if the plane is replaced by any surface applicable to it, say by a parabolic cylinder. But the intrinsic geometry of the sphere essentially differs from that of the plane. For example, the sum of the angles of a spherical triangle is always greater than π.

---

* The considerations concerning the possibility of distinguishing between rectilinear and curvilinear geometric configurations on the basis of studying their intrinsic properties make sense not only for two-dimensional geometric objects, i.e. surfaces, but also for the objects of higher dimensions, in particular for the three-dimensional space. These questions are very important for investigating the general geometric properties of the universe but we cannot dwell in more detail on these problems here.

3. Surfaces of Constant Curvature. Consider a surface whose total curvature K is the same at all its points. Such surfaces are called the **surfaces of constant curvature.** The total curvature being invariant under a bending, it follows that two surfaces of constant curvature are applicable to each other if and only if their total curvatures are equal. It can be shown that, conversely, any two surfaces of the same constant total curvature are always applicable. Hence,

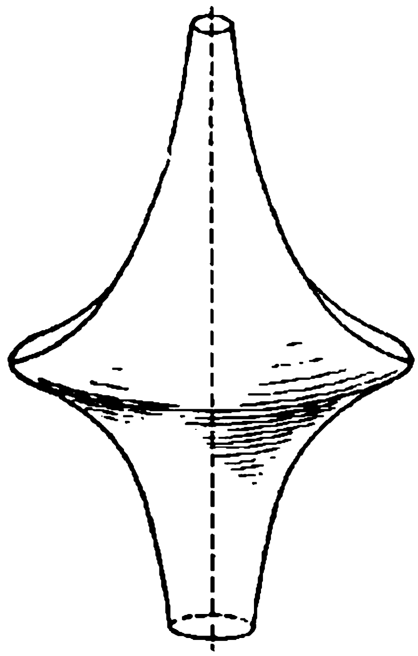


Fig. 3.29

such a surface is completely characterized, from the point of view of its intrinsic geometry, by a single number, that is by its total curvature K.

The geometric properties of surfaces of constant curvature essentially depend on the sign of the curvature, and therefore we must separately consider the surfaces of a positive, negative and zero curvature.

The plane is a surface of zero total curvature. Its intrinsic geometry, as has already been mentioned, is ordinary planimetry. Any other surface of zero total curvature has the same intrinsic geometry.

A sphere of radius R can be regarded as a "canonical model" of a surface with positive constant curvature $K = \frac{1}{R^2}$. The intrinsic geometry of this surface differs from planimetry which is familiar to us. For example, if the arcs of the shortest lengths joining two points (these are the arcs of the great circles in the case of the sphere) are understood as being "straight lines" we can assert that any two "straight lines" intersect when infinitely continued, the sum of the angles of any triangle exceeds π etc.

The so-called **pseudosphere** depicted in Fig. 3.29 is an example of a surface of a negative constant curvature $K < 0$. This is a sur-

face generated by revolution of a **tractrix**, i.e. a plane curve described by the parametric equations

$$x = a\left(\cos t + \ln \tan \frac{t}{2}\right), \qquad y = a \sin t$$

where x and y are Cartesian coordinates, about its asymptote $y = 0$.

The surface, as seen in Fig. 3.29, is not smooth and has a *cuspidal edge*. It is possible to prove that there cannot exist an infinitely continuable smooth surface of constant negative curvature in the three-dimensional space. The intrinsic geometry of the pseudosphere differs from ordinary planimetry and from the intrinsic geometry of the sphere. It coincides with the so-called **Lobachevskian geometry*** in which the sum of the angles of every triangle is less than π, for each point there exists an infinitude of straight lines passing through the point and not intersecting a given straight line etc.

We cannot discuss these problems at length here although they are very important and closely related to modern ideas of physics and, in particular, to the theory of relativity. For these questions we refer the reader to special monographs**.

---

* N. I. Lobachevsky (1792-1856), a famous Russian mathematician, the founder of non-Euclidean geometry.

** E.g. see [18].

4 Line Integrals

Such problems as determining the mass of a material line from its density, computing the work of a field of force along a path and many others require the introduction of the so-called *line integrals* that is the integrals of functions defined over curves. The present chapter is devoted to this notion which is important for mathematical analysis and its applications to physics.

Various physical problems involving integration of functions defined along curves lead to two types of line integrals usually referred to as *line integrals of the first* and *of the second type*. As will be shown, the integrals of these two types can be reduced to each other.

§ 1. LINE INTEGRALS OF THE FIRST TYPE

1. Definition of Line Integral of the First Type. Let AB be a smooth or piecewise smooth* plane curve and $f(M)$ be a function defined on the curve.

Consider a partition of the curve into parts $A_{i-1}A_i$ by means of points of division

$$A = A_0, A_1, \ldots, A_n = B \tag{4.1}$$

and choose an arbitrary point M_i on each arc $A_{i-1}A_i$. Now form the sum

$$\sum_{i=1}^{n} f(M_i)\,\Delta l_i \tag{4.2}$$

where Δl_i is the length of the arc $A_{i-1}A_i$ (Fig. 4.1). We shall refer to such sums as **integral sums**. Let us introduce the following definition.

---

* A curve represented by equations $x = \varphi(t)$, $y = \psi(t)$ is said to be smooth if the functions $\varphi(t)$ and $\psi(t)$ are continuous and possess the continuous derivatives $\varphi'(t)$ and $\psi'(t)$ which do not vanish simultaneously, i.e. if the curve has a tangent at each point and the position of the tangent continuously depends on the point of tangency. A continuous curve composed of a finite number of smooth curves is called piecewise smooth.

Definiton. If integral sums (4.2) *tend to a finite limit** J, *as* max Δl_i *approaches zero, the limit is called the* ***line integral of the first***

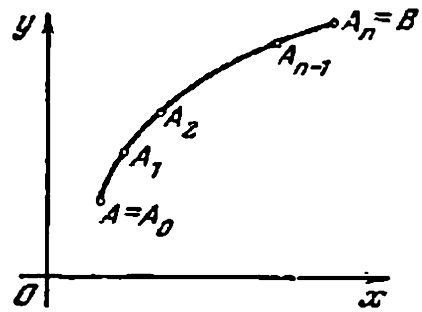


Fig. 4.1

type of the function $f(M)$ ***over the curve*** AB. *We shall denote the integral by the symbol*

$$\int_{AB} f(M)\,dl \tag{4.3}$$

The points of the curve AB being determined by their coordinates (x, y), we shall also designate the function $f(M)$ defined over AB by $f(x, y)$ and write the integral $\int_{AB} f(M)\,dl$ in the form

$$\int_{AB} f(x, y)\,dl$$

But the reader should bear in mind that the variables x and y are not independent because the point (x, y) belongs to the curve AB.

We can easily show that the notion of a line integral of the first type does not in fact essentially differ from that of a definite integral of a function of one independent variable and can be reduced to it. Indeed, let us take the arc length l reckoned from the initial point A as a parameter for the curve AB and write down the equations of the curve in the form

$$x = x(l), \quad y = y(l), \quad 0 \leqslant l \leqslant L \tag{4.4}$$

where L is the length of the entire curve AB. Then an arbitrary function $f(x, y)$ defined on AB reduces to a function $f(x(l), y(l))$ of a single variable l. Let l_i^* be the value of the parameter l corresponding to the point M_i, $i = 1, 2, \ldots, n$. Then integral sum

---

* As in the case of the definite integral (e.g. see [8], Chapter 10, § 1), a number J is said to be the limit of the integral sums if, for every $\varepsilon > 0$, the inequality $\left| J - \sum_{i=1}^{n} f(M_i)\,\Delta l_i \right| < \varepsilon$ holds when max Δl_i becomes sufficiently small.

(4.2) can be rewritten in the form

$$\sum_{i=1}^{n} f(x(l_i^*),\ y(l_i^*))\,\Delta l_i \tag{4.5}$$

which is nothing but an integral sum corresponding to the definite integral

$$\int_0^L f(x(l),\ y(l))\,dl$$

Integral sums (4.2) and (4.5) being equal, the integrals they are related to are also equal. Thus, we have

$$\int_{AB} f(M)\,dl = \int_0^L f(x(l),\ y(l))\,dl \tag{4.6}$$

and both integrals exist or do not exist simultaneously. Consequently, if the function $f(M)$ is continuous* (or piecewise continuous and bounded) along a piecewise smooth curve AB line integral (4.3) is sure to exist because, under these conditions, the definite integral on the right-hand side of equality (4.6) exists.

Note. Although, as has been shown, the line integral of the first type can be directly reduced to the definite integral there is a distinction between the two notions. The matter is that the quantities Δl_i (the lengths of the arcs $A_{i-1}A_i$) are necessarily positive irrespective of which of the end-points A or B, of the curve AB, has been chosen as the initial point. Hence, the orientation of the curve AB (i.e. the choice of a certain direction on it from its initial point to the terminal one) by no means affects the value of integral (4.3) and consequently we have

$$\int_{AB} f(M)\,dl = \int_{BA} f(M)\,dl \tag{4.7}$$

But, as is known, the definite integral $\int_a^b f(x)\,dx$ changes its sign when the limits of integration are interchanged.

When reducing a line integral of the first type to the corresponding definite integral we can as well use any arbitrary parameter t instead of the arc length. Suppose a curve AB is given by parametric equations

$$x = \varphi(t), \quad y = \psi(t) \ (t_0 \leqslant t \leqslant t_1) \tag{4.8}$$

---

* We say that a function $f(M)$ defined on a rectifiable curve is continuous on the curve if it is continuous as a function of the parameter l.

where the functions $\varphi(t)$ and $\psi(t)$ are continuous, their derivatives $\varphi'(t)$ and $\psi'(t)$ are piecewise continuous and bounded and $\varphi'^2(t) + \psi'^2(t) > 0$. Then we can introduce, as a new parameter, the arc length l of the curve AB reckoned from a fixed point. Let us choose the direction in which l is laid off so that the arc length l increases when the parameter t increases. Then l becomes a monotone increasing function of t, and we have

$$dl = \sqrt{\varphi'^2(t) + \psi'^2(t)}\, dt \tag{4.9}$$

Taking advantage of equality (4.6) and formula of changing variable in the definite integral we obtain

$$\int_{AB} f(M)\, dl = \int_0^L f(x(l), y(l))\, dl = \int_{t_0}^{t_1} f(\varphi(t), \psi(t)) \sqrt{\varphi'^2(t) + \psi'^2(t)}\, dt$$

where $t_0 < t_1$. Thus, we arrive at the following theorem.

Theorem 4.1. Let AB be a smooth curve represented by equations

$$x = \varphi(t), \quad y = \psi(t) \quad (t \in [t_0, t_1])$$

and $f(x, y)$ be a function defined along the curve. Then we have the equality

$$\int_{AB} f(x, y)\, dl = \int_{t_0}^{t_1} f(\varphi(t), \psi(t)) \sqrt{\varphi'^2(t) + \psi'^2(t)}\, dt \tag{4.10}$$

provided the integrals entering into it exist; the line integral on the left-hand side exists if and only if the definite integral on the right-hand side exists.

In particular, if the curve AB is determined by an equation of the form

$$y = y(x) \quad (a \leqslant x \leqslant b)$$

expressing y as an explicit function of x formula (4.10) for reducing the line integral to the definite integral takes the form

$$\int_{AB} f(M)\, dl = \int_a^b f(x, y(x)) \sqrt{1 + y'^2}\, dx \tag{4.11}$$

Exercise. Consider the line integral of a function $f(x, y)$ over an arc AB represented by its polar equation

$$r = r(\varphi) \quad (\varphi_1 \leqslant \varphi \leqslant \varphi_2)$$

in the form of a definite integral with respect to φ.

Answer.

$$\int\limits_{AB} f(x, y)\,dl = \int\limits_{\varphi_1}^{\varphi_2} f(r\cos\varphi,\ r\sin\varphi)\sqrt{r^2 + r'^2}\,d\varphi$$

Note. As is well known, the definite integral

$$\int\limits_a^b f(x)\,dx$$

of a nonnegative function can be interpreted as the area of the curvilinear trapezoid (see Fig. 4.2*a*). Similarly, the line integral

$$\int\limits_{AB} f(M)\,dl$$

can be thought of, on condition that $f(M) \geqslant 0$, as the area of piece of a cylindrical surface composed of the perpendiculars to th

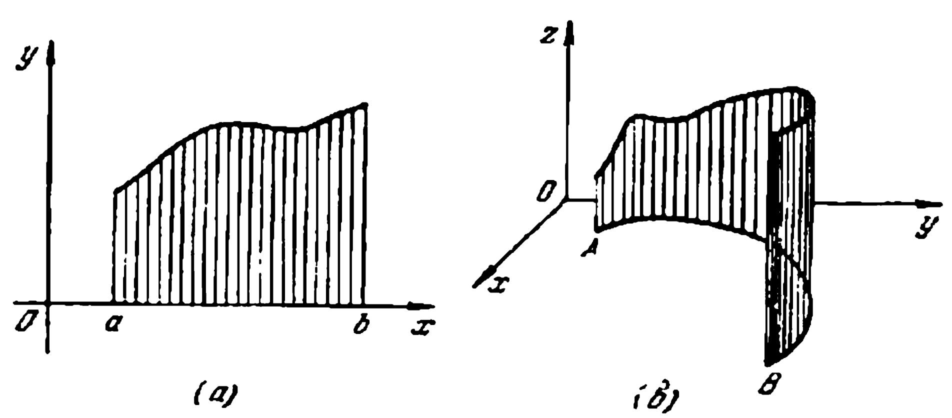


Fig. 4.2

x, y-plane erected at the points of the curve AB and having a variable length $f(M)$ (Fig. 4.2*b*).

2. Properties of Line Integrals. The basic properties of the lin integral of the first type are almost completely analogous to thos of the definite integral and are directly implied by formula (4.6 which reduces the line integral to the definite one. Let us enumerat them.

1 (*linearity*). If $k = \text{const}$ and $f(M)$ is integrable over AB we hav

$$\int\limits_{AB} kf(M)\,dl = k\int\limits_{AB} f(M)\,dl$$

and the integral on the left-hand side exists.

2 (*linearity*). If $f(M)$ and $g(M)$ are integrable on AB the expression $f(M) \pm g(M)$ is also integrable and the relation

$$\int\limits_{AB} (f(M) \pm g(M))\,dl = \int\limits_{AB} f(M)\,dl \pm \int\limits_{AB} g(M)\,dl$$

takes place.

3 (*monotonicity*). If $f(M)$ is a nonnegative integrable function we always have

$$\int_{AB} f(M)\,dl \geqslant 0$$

4 (*additivity*). If an arc AB is composed of two arcs AC and CB the equality

$$\int_{AB} f(M)\,dl = \int_{AC} f(M)\,dl + \int_{CB} f(M)\,dl$$

holds when the integrals entering into it exist, and the integral on the left-hand side exists if and only if both integrals on the right-hand side exist.

5 (*estimation of the modulus of the integral*). If $f(M)$ is integrable over AB the function $|f(M)|$ is also integrable and there is an inequality of the form

$$\left|\int_{AB} f(M)\,dl\right| \leqslant \int_{AB} |f(M)|\,dl$$

6 (*mean value theorem*). If $f(M)$ is continuous on a curve AB there is a point M^* belonging to the curve such that

$$\int_{AB} f(M)\,dl = f(M^*)\,L$$

where L is the length of the arc AB.

7 (*independence of the line integral of the first type of the orientation of the path of integration*). As has been mentioned, we always have

$$\int_{AB} f(M)\,dl = \int_{BA} f(M)\,dl$$

and hence the choice of the direction of the arc AB does not affect the value of the integral of an arbitrary scalar function $f(M)$ taken along the arc.

3. Some Applications of Line Integrals of the First Type. We shall point out some typical problems whose solution naturally involves line integrals of the first type.

(1) *Determining the mass of a material line from its density.* A material line will be understood as a piecewise smooth curve along which a mass is distributed. The linear density $\rho(M)$ of a material line at a point M is the limit to which the ratio of the mass $\Delta\mu$, carried by an arc MM', to the length of the arc MM' tends on condition that the arc length tends to zero. In other words, if l is the length of an arc AM and $\mu(M)$ is the mass of the arc we have $\rho(M) = \frac{d\mu(l)}{dl}$. It follows that the mass μ_{AB} of the arc AB is expressed

by the integral $\int\limits_0^l \rho\, dl$, that is by the line integral

$$\int\limits_{AB} \rho(M)\, dl$$

of the density ρ taken along the curve AB.

(2) *Finding the coordinates of the centre of gravity of a material line.* Let a mass be distributed, with density $\rho(x, y)$,* along a curve AB. After the curve has been broken up into parts of lengths Δl_i, $i = 1, \ldots, n$, and an arbitrary point (x_i, y_i) has been taken on each part we can approximately represent the material line as the system of mass points $\rho(x_i, y_i)\Delta l_i$ placed at the points (x_i, y_i). The centre of gravity of such a system of material points has the coordinates

$$x_c = \frac{\sum\limits_{i=1}^{n} x_i \rho(x_i, y_i)\, \Delta l_i}{\sum\limits_{i=1}^{n} \rho(x_i, y_i)\, \Delta l_i}, \quad y_c = \frac{\sum\limits_{i=1}^{n} y_i \rho(x_i, y_i)\, \Delta l_i}{\sum\limits_{i=1}^{n} \rho(x_i, y_i)\, \Delta l_i}$$

These expressions can be regarded as approximate values of the coordinates x_c and y_c of the centre of gravity of the material line AB. To obtain the exact values of the coordinates we must pass to the limit as $\max \Delta l_i \to 0$. The passage to the limit results in

$$x_c = \frac{\int\limits_{AB} x\rho(x, y)\, dl}{\int\limits_{AB} \rho(x, y)\, dl}, \quad y_c = \frac{\int\limits_{AB} y\rho(x, y)\, dl}{\int\limits_{AB} \rho(x, y)\, dl} \tag{4.12}$$

In particular, for a homogeneous material line with $\rho = \text{const}$ we obtain

$$x_c = \frac{\int\limits_{AB} x\, dl}{\int\limits_{AB} dl}, \quad y_c = \frac{\int\limits_{AB} y\, dl}{\int\limits_{AB} dl} \tag{4.13}$$

(3) *Computing moments of inertia of a material line.* The moment of inertia of a system of mass points m_i $(i = 1, \ldots, n)$ about a straight line is equal to

$$\sum_{i=1}^{n} r_i^2 m_i$$

where r_i is the distance from the ith mass to the straight line. In particular, for a system of material points lying in the x, y-plane

* Here and henceforth it appears natural to determine the points of a curve by their Cartesian coordinates x and y (see Sec. 1).

the moments of inertia about the x-axis and the y-axis are respectively equal to

$$I_x=\sum_{i=1}^{n} y_i^2 m_i \quad \text{and} \quad I_y=\sum_{i=1}^{n} x_i^2 m_i$$

where (x_i, y_i) are the coordinates of the mass point m_i. To get the moments of inertia about the coordinate axes of a material line AB along which a mass is distributed with density $\rho(x, y)$ we must perform the passage to the limit similar to the one in the preceding problem. Then, for the moments of inertia of the curve AB about the coordinate axes, we deduce the expressions

$$I_x=\int_{AB} y^2\rho(x, y)\,dl, \quad I_y=\int_{AB} x^2\rho(x, y)\,dl \tag{4.14}$$

(4) *Gravitational attraction of a mass point by a material line.* Let AB be a material line with mass density $\rho(x, y)$. We now consider the expressions of the projections on the coordinate axes of the gravitational force $\mathbf{F}$ with which the material line attracts a mass point m_0 having the coordinates (x_0, y_0). Applying the argument analogous to the one presented above we find that the projections F_x and F_y are given by the formulas

$$F_x=\gamma m_0\int_{AB}\frac{\rho(x, y)(x-x_0)}{r^3}\,dl, \quad F_y=\gamma m_0\int_{AB}\frac{\rho(x, y)(y-y_0)}{r^3}\,dl$$

where γ is the constant of gravitation and $r=\sqrt{(x-x_0)^2+(y-y_0)^2}$.

If integration of a vector is understood in the sense that each projection is integrated (see § 1 in Chapter 3) we can replace the two scalar formulas by a single vector relation and thus the force $\mathbf{F}$ with which the material point m_0 is attracted by the material line AB is equal to

$$\mathbf{F}=\gamma m_0\int_{AB}\frac{\rho(x, y)}{r^3}\,\mathbf{r}\,dl \tag{4.15}$$

where $\mathbf{r}$ is the vector whose projections are $(x-x_0)$ and $(y-y_0)$.

4. Line Integrals of the First Type in Space. The definition of a line integral of the first type over a plane curve is directly transferred to the case of a function $f(M)$ defined along a space curve.

If a space curve is represented by parametric equations

$$x=\varphi(t), \quad y=\psi(t), \quad z=\chi(t) \quad (t_0\leqslant t\leqslant t_1)$$

the line integral of the first type of the function $f(M)$ taken along the curve is reduced to the definite integral by means of the formula

$$\int_{AB} f(M)\,dl=\int_{t_0}^{t_1} f(\varphi(t), \psi(t), \chi(t))\sqrt{\varphi'^2(t)+\psi'^2(t)+\chi'^2(t)}\,dt$$

The conditions for the existence and basic properties of line integrals in space are completely analogous to those for the case of a plane. The line integrals of the first type in space are naturally encountered in such problems as determining the mass of a space curve (material line) from a given density, finding the coordinates of the centre of gravity and the moments of inertia of the curve and the like. By analogy with the arguments presented for the case of a plane curve, the reader can easily write the corresponding formulas.

§ 2. LINE INTEGRALS OF THE SECOND TYPE

1. Statement of the Problem. Work of a Field of Force. Now we introduce integrals of another kind, the so-called *line integrals of the second type*.

To begin with we shall take a concrete physical problem. Consider a plane *field of force*, that is a domain in a plane at whose each point M a force $\mathbf{F}(M)$ is determined. The projections of the force on the x-axis and y-axis will be denoted, respectively, as $P(M)$ and $Q(M)$.

Let us find the work of the force field when a point moves along a curve AB.

If the force $\mathbf{F}$ is constant (both in its absolute value and direction) and the path AB is rectilinear the corresponding work is equal to the product of the value of the force by the path length and by the cosine of the angle between the force and the direction of displacement, i.e. to the scalar product

$$(\mathbf{F}, \overline{AB})$$

Let us now find the expression of the work in the general case when the force $\mathbf{F}$ may be variable and the trajectory of motion curvilinear. Let AB be a smooth curve lying in the domain where the field of force is defined. Divide the curve AB into parts by means of points

$$A = M_0, M_1, \ldots, M_n = B$$

and consider the broken line with vertices at the points M_i (see Fig. 4.3). We can approximately consider the force $\mathbf{F}$ to retain a constant value along each segment $M_{i-1}M_i$ of the broken line, say equal to $F(M_i)$, and compute the work corresponding to the motion along the broken line. If (x_i, y_i) are the coordinates of the point M_i and

$$\Delta x_i = x_i - x_{i-1}, \quad \Delta y_i = y_i - y_{i-1}$$

the work corresponding to the displacement along the segment $M_{i-1}M_i$ is equal to

$$(\mathbf{F}(M_i), \overline{M_{i-1}M_i}) = P(M_i)\,\Delta x_i + Q(M_i)\,\Delta y_i$$

and the total work corresponding to the motion along the whole broken line is equal to

$$\sum_{i=1}^{n} (P(M_i)\,\Delta x_i + Q(M_i)\,\Delta y_i) \qquad (4.16)$$

The last sum can be approximately taken as the expression of the work performed by the field of force **F** (M) along the curve AB. To derive the exact value of the work we must pass to the limit in sum

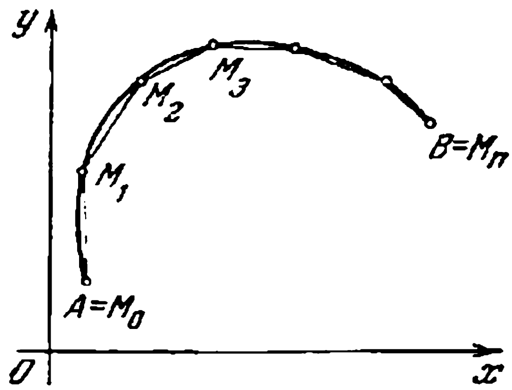


Fig. 4.3

(4.16) by making the maximum of the lengths of the arcs $M_{i-1}M_i$ tend to zero. We now consider such a passage to the limit in the general form.

2. Definition of Line Integral of the Second Type. Let AB be a smooth curve and $\mathbf{F}(M) = (P(M), Q(M))$ be a vector function defined on the curve AB. Break up the curve into parts by means of points

$$A = M_0,\ M_1,\ \ldots,\ M_n = B$$

whose coordinates are, respectively $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$. Take the sum

$$T = \sum_{i=1}^{n} [P(M_i)\,\Delta x_i + Q(M_i)\,\Delta y_i] \qquad (4.17)$$

where $\Delta x_i = x_i - x_{i-1}$ and $\Delta y_i = y_i - y_{i-1}$.

If all such sums tend to a certain finite limit when the maximum of the lengths of the arcs $M_{i-1}M_i$ tends to zero the limit is called the **line integral of the second type of the vector function** $\mathbf{F} = (P, Q)$ and is denoted* by the symbol

$$\int_{AB} P(M)\,dx + Q(M)\,dy \qquad (4.18)$$

* We shall sometimes write $P(x, y)$ and $Q(x, y)$ instead of $P(M)$ and $Q(M)$ where x and y should be understood as Cartesian coordinates of the moving point M. We shall also simply denote the functions $P(M)$ and $Q(M)$ by P and Q and write line integral (4.18) in the form $\int_{AB} P\,dx + Q\,dy$ or $\int_{A}^{B} P\,dx + Q\,dy$ unless this leads to misunderstanding.

The integral is obviously the sum of the two integrals

$$\int_{AB} P(M)\,dx \quad \text{and} \quad \int_{AB} Q(M)\,dy$$

corresponding to the vectors $(P, 0)$ and $(0, Q)$, the components into which the vector (P, Q) is resolved.

Note. The line integral of the second type must not be confused with an integral of a vector with respect to a scalar parameter (which has been encountered in § 1, Sec. 5 of Chapter 3 and at the end of Sec. 5, § 1 of the present chapter) when the projections of the vector are integrated separately (e.g. we have had such a situation in the computation of the force of attraction of a material point by a material line).

3. Connection Between Line Integrals of the First and the Second Types. The line integral of the second type can be easily reduced to the line integral of the first type considered in § 1. The relationship between the integrals is described by the following theorem:

Theorem 4.2. Let AB be a smooth curve determined by equations

$$x = x(l), \qquad y = y(l) \tag{4.19}$$

and $\mathbf{F} = (P, Q)$ *be a vector function defined on the curve and bounded* on it. Then we have the equality*

$$\int_{AB} P\,dx + Q\,dy = \int_{AB} (P\cos\alpha + Q\sin\alpha)\,dl \tag{4.20}$$

where $\alpha = \alpha(M)$ *is the angle between the tangent to the curve AB at the point M and the positive direction of the x-axis, provided that the integrals entering into (4.20) exist. Furthermore, the integral on the left-hand side exists if and only if the line integral of the first type on the right-hand side of relation (4.20) exists.*

Proof. Let us prove the equality

$$\int_{AB} P\,dx = \int_{AB} P\cos\alpha\,dl$$

The equality

$$\int_{AB} Q\,dy = \int_{AB} Q\sin\alpha\,dl$$

is proved similarly.

---

* A vector function (P, Q) is said to be bounded if the scalar functions P and Q are bounded.

The integral

$$\int_{AB} P\,dx$$

is, by definition, the limit of the sums of the form

$$T=\sum_{i=1}^{n} P(M_i)\,\Delta x_i$$

Let us compare the sum with the integral sum

$$T^*=\sum_{i=1}^{n} P(M_i)\cos\alpha(M_i)\,\Delta l_i$$

associated (for the same partition of the curve AB) with the integral

$$\int_{AB} P\cos\alpha\,dl$$

If $x=x(l)$ the relation

$$\frac{dx}{dl}=\cos\alpha(M) \tag{4.21}$$

holds for each point M of the curve AB and, consequently, we have

$$\Delta x_i=\int_{l_{i-1}}^{l_i}\cos\alpha\,dl$$

Applying the mean value theorem we thus obtain

$$\Delta x_i=\cos\alpha(M_i^*)\,\Delta l_i$$

where M_i^* is a point belonging to the arc $M_{i-1}M_i$. Therefore,

$$|T-T^*|=\left|\sum_{i=1}^{n} P(M_i)\,[\cos\alpha(M_i)-\cos\alpha(M_i^*)]\,\Delta l_i\right|\leqslant$$

$$\leqslant\sum_{i=1}^{n}|P(M_i)|\cdot|\cos\alpha(M_i)-\cos\alpha(M_i^*)|\,\Delta l_i$$

The function $\cos\alpha(M)$ is continuous on the smooth curve AL and hence it is uniformly continuous because the curve is a bounded closed set of points. Thus, given any $\varepsilon>0$, the inequality

$$|\cos\alpha(M_i)-\cos\alpha(M_i^*)|<\varepsilon$$

is fulfilled for each partition of the curve AB whose fineness is small enough. This implies

$$|T-T^*|\leqslant C\varepsilon\sum_{i=1}^{n}\Delta l_i=C\varepsilon L$$

where $C = \sup |P|$ and L is the length of the curve AB. It follows that if the integral sums T^* have a limit the sums T tend to the same limit. The theorem has thus been proved.

Note. The expression $P\cos\alpha + Q\sin\alpha$ is the scalar product $(\mathbf{F}, \boldsymbol{\tau})$ of the vector $\mathbf{F} = (P, Q)$ by the unit vector $\boldsymbol{\tau} = (\cos\alpha, \sin\alpha)$ tangent to the curve AB and hence is equal to the projection of the vector $\mathbf{F} = (P, Q)$ on the tangent line to AB. Denoting the projection by the symbol F_τ and taking advantage of equality (4.20) we can write line integral (4.18) in the form

$$\int_{AB} F_\tau \, dl \tag{4.22}$$

This abridged notation will often be used in what follows and particularly in Chapter 6. The integral is sometimes also written in the form

$$\int_{AB} (\mathbf{F}, d\mathbf{l}) \tag{4.23}$$

where $d\mathbf{l}$ is understood as an infinitesimal vector whose projections on the coordinate axes are

$$dx = dl\cos\alpha \quad \text{and} \quad dy = dl\sin\alpha$$

4. Evaluating Line Integral of the Second Type. Theorems 4.1 and 4.2 immediately imply the following result.

Theorem 4.3. Let AB be a smooth curve represented by equations

$$x = \varphi(t), \qquad y = \psi(t) \tag{4.24}$$

and let $\mathbf{F} = (P, Q)$ be a vector function defined on the curve. Then we have the relation

$$\int_{AB} P\,dx + Q\,dy = \int_{t_0}^{t_1} [P(\varphi(t), \psi(t))\varphi'(t) + Q(\varphi(t), \psi(t))\psi'(t)]\,dt \tag{4.25}$$

provided that the integrals entering into it exist, and the integral on the left-hand side is sure to exist if the definite integral on the right-hand side exists. Here, the value t_0 of the parameter t corresponds to the point A and t_1 to the point B.

Theorems 4.1-4.3 obviously remain true if the curve AB is not smooth but piecewise smooth.

Let us consider some important special cases of formula (4.25).

If the curve AB is determined by an equation

$$y = y(x) \tag{4.26}$$

expressing y as an explicit function of the variable x which runs over an interval $[a, b]$ formula (4.25) for reducing the line integral of the second type to the definite integral turns into

$$\int_{AB} P\,dx + Q\,dy = \int_a^b [P(x, y(x)) + Q(x, y(x))\,y'(x)]\,dx \quad (4.27)$$

where the value $x = a$ corresponds to the initial point A of the curve AB and $x = b$ to its terminal point B. In particular, if the curve AB is a segment of a horizontal straight line $y = y_0$ we have $y' \equiv 0$ along it and thus the integral

$$\int_{AB} P\,dx + Q\,dy$$

taken over such a segment simply reduces to the integral

$$\int_a^b P(x, y_0)\,dx$$

Similarly, for a curve determined by an equation

$$x = x(y) \quad (4.28)$$

where y ranges in an interval $[c, d]$ we obtain

$$\int_{AB} P\,dx + Q\,dy = \int_c^d [P(x(y), y)\,x'(y) + Q(x(y), y)]\,dy \quad (4.29)$$

If AB is a segment of a vertical straight line $x = x_0$ we have $x' \equiv 0$ and integral (4.29) takes the form

$$\int_{AB} Q(x_0, y)\,dy \quad (4.30)$$

Examples

1. Evaluate the integral

$$\int_{AB} x^2\,dx + xy\,dy \quad (4.31)$$

taken along

(a) the line segment drawn from the point (1, 0) to the point (0, 1),

(b) the quarter circle $x = \cos t$, $y = \sin t$ $\left(0 \leqslant t \leqslant \frac{\pi}{2}\right)$ joining the same points (see Fig. 4.4).

Solution.

(a) $$\int\limits_{AB} x^2\,dx + xy\,dy = \int\limits_1^0 (x^2 - x(1-x))\,dx = \int\limits_1^0 (2x^2 - x)\,dx = -\frac{1}{6}$$

(b) $$\int\limits_{AB} x^2\,dx + xy\,dy = \int\limits_0^{\frac{\pi}{2}} (-\cos^2 t \sin t + \cos^2 t \sin t)\,dt = 0$$

2. Evaluate the integral

$$\int\limits_{AB} 3x^2y\,dx + (x^3+1)\,dy \tag{4.32}$$

taken along

(a) the line segment drawn from the point (0, 0) to the point (1, 1),

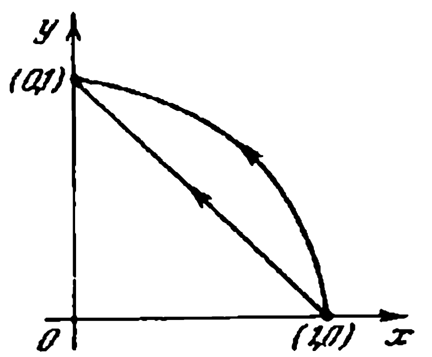


Fig. 4.4

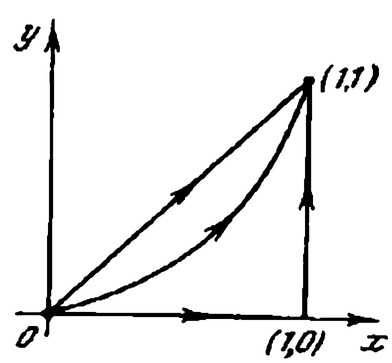


Fig. 4.5

(b) the arc of the parabola $y = x^2$ connecting the same points,

(c) the broken line passing through the points (0, 0), (1, 0) and (1, 1) (Fig. 4.5).

Solution.

(a) $$\int\limits_{AB} 3x^2y\,dx + (x^3+1)\,dy = \int\limits_0^1 (4x^3+1)\,dx = 2$$

(b) $$\int\limits_{AB} 3x^2y\,dx + (x^3+1)\,dy = \int\limits_0^1 (5x^4+2x)\,dx = 2$$

(c) $$\int\limits_{AB} 3x^2y\,dx + (x^3+1)\,dy =$$

$$= \int\limits_{(0,\,0)}^{(1,\,0)} 3x^2y\,dx + \int\limits_{(1,\,0)}^{(1,\,1)} (x^3+1)\,dy = \int\limits_0^1 2\,dy = 2$$

Note. The reader has undoubtedly noticed that the integral in the second example assumes the same value when taken along the three different paths connecting the two given points. This fact is connected with an important property of the line integrals of the second type which will be elucidated in § 4.

5. Dependence of Line Integral of the Second Type on the Orientation of the Path of Integration. The definition of the line integral

$$\int\limits_{AB} P\,dx + Q\,dy \tag{4.33}$$

automatically suggests that a constant factor can be taken outside the integral sign, the integral of a sum of two vector functions is equal to the sum of the integrals of the addends etc.

Another important property of integral (4.33) should be emphasized here: the line integral of the second type, in contrast to the line integral of the first type, defined in § 1, depends on the orientation of the curve AB it is taken over, namely, it changes its sign when the orientation of the curve is reversed:

$$\int\limits_{BA} P\,dx + Q\,dy = -\int\limits_{AB} P\,dx + Q\,dy \tag{4.34}$$

For, if we change the direction in which the curve AB is traversed we thus replace the quantities Δx_i and Δy_i entering into sum (4.17) by $-\Delta x_i$ and $-\Delta y_i$. This changes the sign of integral sums (4.17) and, consequently, the sign of their limit as well.

This property of the line integral of the second type is coherent with the physical interpretation of the integral as the work of a field of force along a path. Indeed, if the direction of tracing the trajectory is reversed the work performed by the force field changes its sign to the opposite.

6. Line Integrals Along Self-Intersecting and Closed Paths. For the applications of the theory of line integrals, it is advisable to include in our consideration the paths of integration which may have points of self-intersection because such cases are encountered in various problems. Mathematically, this means that when we consider a curve determined by equations

$$x = x(t), \quad y = y(t) \quad (a \leqslant t \leqslant b)$$

we should not exclude the cases in which there may exist two (or more) distinct values t_1 and t_2 of the parameter t such that

$$x(t_1) = x(t_2) \quad \text{and} \quad y(t_1) = y(t_2)$$

When studying such cases for the line integral of the second type we must take into account that the specification of a path of integration includes the indication of a set of points constituting the curve in question as well as a certain direction in which the path is described. If a curve has points of self-intersection the way it is traversed is not completely characterized by setting initial and

terminal points. For instance, the curves shown in Fig. 4.6*a* and *b* should be regarded as being two different paths although they coincide as sets of points. What has been stipulated here pertains not only to the plane curves but also to the space ones.

We also deal with line integrals taken over various closed contours. A closed contour (in the plane) is understood as a curve

$$x = x(t), \quad y = y(t), \quad (a \leqslant t \leqslant b)$$

such that

$$x(a) = x(b) \quad \text{and} \quad y(a) = y(b)$$

Here we can also have the cases when such a contour intersects itself, that is there may exist various values of the parameter t, other than $t = a$ and $t = b$, for which the corresponding values of x and y, respectively, coincide.

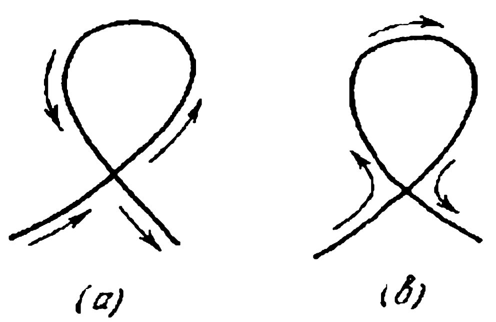


Fig. 4.6

If a closed contour, in the plane, has no points of self-intersection there are only two possible directions of describing it: the contour can be traced counterclockwise (the positive orientation) or clockwise (the negative orientation).

If an integral of the second type

$$\int_C P\,dx + Q\,dy$$

is taken along such a contour its values corresponding to the two different orientations of the contour C are equal in their absolute values and have the opposite signs. In what follows, we shall, as a rule, consider closed contours with the positive orientation and replace line integrals of the second type taken along negatively oriented contours by the corresponding integrals taken in the positive direction with the minus sign attached to them.

A line integral over a closed contour is often denoted by the symbol

$$\oint_C P\,dx + Q\,dy$$

7. Line Integral of the Second Type Over a Space Curve. We have considered line integrals, of vector functions, taken along plane curves. Most of the facts of the theory are automatically generalized to the case of a space curve.

Let AB be a smooth space curve and $\mathbf{F} = (P, Q, R)$ be a continuous vector function defined on the curve. We divide the curve AB into parts by means of points of division

$$A = M_0, M_1, \ldots, M_n = B$$

having coordinates (x_i, y_i, z_i), $i = 1, 2, \ldots, n$, and form the sum

$$\sum_{i=1}^{n} \{P(M_i)\,\Delta x_i + Q(M_i)\,\Delta y_i + R(M_i)\,\Delta z_i\}$$

where

$$\Delta x_i = x_i - x_{i-1}, \quad \Delta y_i = y_i - y_{i-1} \quad \text{and} \quad \Delta z_i = z_i - z_{i-1}$$

The limit of such sums will be referred to as the **line integral of the second type of the vector function $\mathbf{F} = (P, Q, R)$ over the space curve** AB and denoted by

$$\int\limits_{AB} P(M)\,dx + Q(M)\,dy + R(M)\,dz \tag{4.35}$$

or*

$$\int\limits_{AB} P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz$$

Applying arguments completely similar to those used in investigating the case of a plane curve we arrive at the following formula reducing integral (4.35) to the line integral of the first type:

$$\int\limits_{AB} P\,dx + Q\,dy + R\,dz = \int\limits_{AB} [P\cos\alpha + Q\cos\beta + R\cos\gamma]\,dl$$

where α, β and γ are the angles formed by the tangent line to the curve AB (drawn in the direction of tracing the path of integration) with the positive directions of the coordinate axes x, y and z.

If a smooth curve AB is given by equations

$$x = \varphi(t), \quad y = \psi(t), \quad z = \chi(t)$$

---

* For brevity, we shall sometimes write the integral in the form $\int\limits_{AB} P\,dx + Q\,dy + R\,dz$ or $\int\limits_{A}^{B} P\,dx + Q\,dy + R\,dz$ unless this leads to misunderstanding.

and the value $t = t_0$ of the parameter t corresponds to the point A and $t = t_1$ to the point B we have the relation

$$\int_{AB} P\,dx + Q\,dy + R\,dz = \int_{t_0}^{t_1} [P(\varphi(t), \psi(t), \chi(t))\,\varphi'(t) +$$
$$+ Q(\varphi(t), \psi(t), \chi(t))\,\psi'(t) + R(\varphi(t), \psi(t), \chi(t))\,\chi'(t)]\,dt \quad (4.36)$$

which reduces the line integral of the second type to the definite integral.

The expression $P\cos\alpha + Q\cos\beta + R\cos\gamma$ being the projection of the vector $\mathbf{F} = (P, Q, R)$ on the tangent line to the curve AB, we can write line integral (4.35), as in the case of a plane curve, in the form

$$\int_{\overset{\frown}{AB}} F_\tau\,dl$$

where F_τ designates the projection of $\mathbf{F}$ on the tangent. All the properties of the line integrals in the plane are automatically transferred to the spatial case. In particular, line integral (4.35) changes its sign when the orientation of the path of integration is replaced by the opposite one, i.e.

$$\int_{\overset{\frown}{BA}} P\,dx + Q\,dy + R\,dz = -\int_{\overset{\frown}{AB}} P\,dx + Q\,dy + R\,dz$$

Accordingly, in the definite integral entering into the right-hand side of formula (4.36), the lower limit of integration t_0 should be understood as the value of the parameter t corresponding to the initial point A of the curve AB and the upper limit t_1 as the one corresponding to the terminal point B irrespective of which of the numbers t_0 and t_1 is greater.

§ 3. GREEN'S FORMULA

Here we shall establish so-called *Green's* formula* which expresses a relationship between a line integral

$$\oint_C P\,dx + Q\,dy$$

taken along the boundary of a domain and a double integral over the domain. The formula is widely used in mathematical analysis as well as in its applications. Some of the applications will be considered later.

---

* Green, George (1793-1841), an English mathematician known for his results in the field of mathematical physics.

1. Derivation of Green's Formula. Let us first take a domain G of a particularly simple form, bounded by piecewise smooth curves

$$y = y_1(x), \quad y = y_2(x), \quad (y_2(x) \geqslant y_1(x)) \tag{4.37}$$

and by vertical line segments

$$x = a, \quad x = b \quad (b > a) \tag{4.38}$$

(see Fig. 4.7). We shall consider the boundary $ABCDA$ of the domain as being positively oriented in the sense that when describing the contour the domain is kept always on the left (in the particular

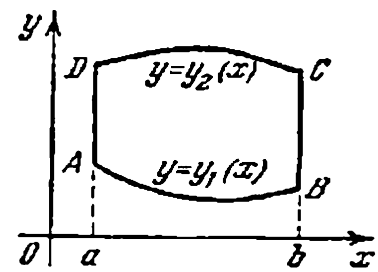


Fig. 4.7

case of the domain G this means that the contour is traced counterclockwise). Let a function $P(x, y)$ be defined and continuous throughout the domain G including its boundary and possess the continuous partial derivative $\frac{\partial P}{\partial y}$ in this region.

Let us consider the double integral $\iint\limits_G \frac{\partial P}{\partial y}\,dx\,dy$ which we shall try to transform into a line integral. To this end we reduce it to the iterated integral $\int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} \frac{\partial P}{\partial y}\,dy$ and perform the integration with respect to y. This yields

$$\iint\limits_G \frac{\partial P}{\partial y}\,dx\,dy = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} \frac{\partial P}{\partial y}\,dy =$$

$$= \int\limits_a^b [P(x, y_2(x)) - P(x, y_1(x))]\,dx =$$

$$= \int\limits_a^b P(x, y_2(x))\,dx - \int\limits_a^b P(x, y_1(x))\,dx \tag{4.39}$$

Each of the last definite integrals can be regarded as a line integral taken along the corresponding arc (see formula (4.27)), namely as

$$\int\limits_a^b P(x, y_2(x))\,dx = \int\limits_{DC} P(x, y)\,dx = -\int\limits_{CD} P(x, y)\,dx$$

and

$$-\int_a^b P(x, y_1(x))\, dx = -\int_{AB} P(x, y)\, dx$$

Now adding the two line integrals

$$-\int_{BC} P(x, y)\, dx \quad \text{and} \quad -\int_{DA} P(x, y)\, dx$$

to the right-hand side of equality (4.39) and taking into account that both integrals are equal to zero (since $dx = 0$ on the vertical line segments) we arrive at the relation

$$\iint_G \frac{\partial P}{\partial y}\, dx\, dy = -\int_{AB} P\, dx - \int_{BC} P\, dx - \int_{CD} P\, dx - \int_{DA} P\, dx$$

that is

$$\iint_G \frac{\partial P}{\partial y}\, dx\, dy = -\int_{ABCDA} P\, dx \tag{4.40}$$

Equality (4.40) has been proved for a domain of a special form bounded by the curves of types (4.37) and (4.38). But the formula can also be extended to an arbitrary domain which can be divided into a finite number of parts of this particular form.

For, if a domain G with a boundary L is broken up into parts G_i, $i = 1, 2, \ldots, n$, such that, for each part, the relation

$$\iint_G \frac{\partial P}{\partial y}\, dx\, dy = -\int_{L_i} P\, dx$$

holds (where L_i is the boundary of G_i) we can sum up these relations with respect to i from 1 to n and thus obtain the double integral taken over the whole domain G on the left-hand side and the sum of the line integrals along the contours L_i on the right-hand side. But each contour L_i consists of some arcs of the boundary of the domain G and of some arcs of the auxiliary curves which divide the domain G into parts. Every arc of each auxiliary curve being a constituent of exactly two contours of the type L_i, we thus see that the line integral over such an arc is taken twice, the directions of integration being opposite (see Fig. 4.8). Therefore, when we add together integrals of the form

$$\int_{L_i} P\, dx$$

the integrals taken over all the auxiliary arcs mutually cancel out and hence only the integral along the boundary of the domain G

remains on the right-hand side. Thus, we obtain the equality

$$\iint_G \frac{\partial P}{\partial y}\,dx\,dy = -\int_L P\,dx \tag{4.41}$$

where L is the positively oriented* boundary of the domain G.

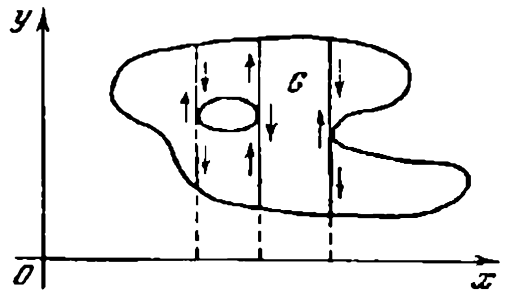


Fig. 4.8

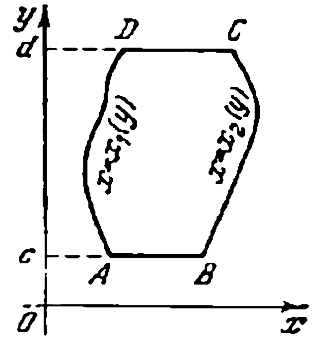


Fig. 4.9

Let us now interchange the roles of the variables x and y and take a domain G bounded by the horizontal line segments

$$y = c, \quad y = d \tag{4.42}$$

and by the curves

$$x = x_1(y), \quad x = x_2(y) \tag{4.43}$$

shown in Fig. 4.9. Let the function $Q(x, y)$ and its derivative $\frac{\partial Q}{\partial x}$ be defined and continuous in the domain G (including its boundary). Taking the double integral

$$\iint_G \frac{\partial Q}{\partial x}\,dx\,dy$$

in the form

$$\int_c^d dy \int_{x_1(y)}^{x_2(y)} \frac{\partial Q}{\partial x}\,dx$$

and performing the same calculations as in deducing formula (4.40) we receive the equality

$$\iint_G \frac{\partial Q}{\partial x}\,dx\,dy = \int_{ABCDA} Q\,dy$$

analogous to (4.40) (with the only distinction that there is no minus sign on the right-hand side). Further, following the arguments

* As has been mentioned, this means that the domain G is always kept on the left of a person walking round the contour L in the chosen (positive) direction of tracing the boundary of G.

similar to the ones given above we find that the relation

$$\iint_G \frac{\partial Q}{\partial x}\,dx\,dy = \int_L Q\,dy \qquad (4.44)$$

is true not only for the domains bounded by the curves of types (4.42) and (4.43) but also for the unions of a finite number of such domains.

A domain G which can be divided into parts with boundaries of form (4.37), (4.38) and of form (4.42), (4.43) will be referred to, for brevity, as a simple domain. As has been shown, for a simple domain, relations (4.41) and (4.44) are fulfilled. Subtracting (4.41) from (4.44) we derive the formula

$$\int_L P\,dx + Q\,dy = \iint_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy \qquad (4.45)$$

where the line integral is taken along the boundary L of the domain G in the positive direction. This very formula, which we set out to prove, is called **Green's formula**. Thus, we can state the following result.

Theorem 4.4. *Let G be a simple domain and functions $P(x, y)$ and $Q(x, y)$ and their derivatives be defined and continuous on the closure of the domain G (i.e. on the union of G and its boundary). Then Green's formula (4.45) holds.*

Note 1. If the boundary L of a domain G is composed of a finite number of separate contours the symbol $\int_L P\,dx + Q\,dy$ should be understood as the sum of the integrals taken over all the contours

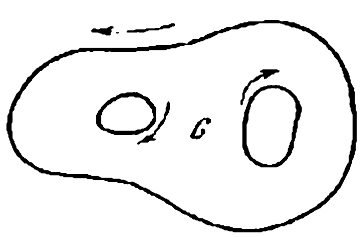


Fig. 4.10

the entire contour is formed of, and each contour is described in the direction such that the domain G always remains on the left of the contour (see Fig. 4.10).

Note 2. In deducing Green's formula we have supposed that the functions P and Q and their partial derivatives $\frac{\partial P}{\partial y}$ and $\frac{\partial Q}{\partial x}$ are continuous not only in the interior of the domain G but also on its boundary. But it turns out that it is sufficient to impose the condi-

tion that the derivatives $\frac{\partial P}{\partial y}$ and $\frac{\partial Q}{\partial x}$ are continuous and bounded in the interior of the domain G. Indeed, take again a domain G bounded by curves $y = y_1(x)$, $y = y_2(x)$ and vertical line segments $x = a$, $x = b$ (Fig. 4.7). Choose a sufficiently small number $\delta > 0$ and consider the domain G_δ bounded above and below by the curves $y = y_2(x) - \delta$ and $y = y_1(x) + \delta$, respectively, and on the sides by the vertical line segments $x = a + \delta$ and $x = b - \delta$. For every sufficiently small $\delta > 0$ such a domain G_δ exists and is strictly contained, together with its boundary, within G, and hence all the conditions for which the validity of equality (4.41) has been established are fulfilled for G_δ. Therefore,

$$\iint_{G_\delta} \frac{\partial P}{\partial y}\,dx\,dy = -\int_{L_\delta} P\,dx \tag{4.46}$$

where L_δ is the boundary of the domain G_δ. The difference between the area of the domain G_δ and that of the domain G not exceeding the quantity $l\delta$ where l is the length of the boundary L of the domain G, the integral on the left-hand side of relation (4.46) differs from

$$\iint_G \frac{\partial P}{\partial y}\,dx\,dy$$

not more than by the quantity $l\delta M$ where M is the supremum of $\left|\frac{\partial P}{\partial y}\right|$ inside G. Furthermore, the function $P(x, y)$ is contin ous and therefore uniformly continuous and bounded in the closure of the domain G. It follows immediately that

$$\int_{L_\delta} P\,dx \to \int_L P\,dx \qquad \text{as } \delta \to 0$$

Hence, we can pass to the limit in equality (4.46), for $\delta \to 0$, and thus prove that the relation

$$\iint_G \frac{\partial P}{\partial y}\,dx\,dy = -\int_L P\,dx$$

is true for the domain shown in Fig. 4.7 and, consequently, it applies to any simple domain. The validity of the equality

$$\iint_G \frac{\partial Q}{\partial x}\,dx\,dy = \int_L Q\,dy$$

is established in a similar way.

It should be noted that the requirement that the derivatives $\frac{\partial P}{\partial y}$ and $\frac{\partial Q}{\partial x}$ are bounded in the domain G can be replaced by the

condition that the integral $\iint\limits_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy$ (understood as a improper double integral in case the derivatives are unbounded exists.

Note 3. Green's formula has been proved here for the domai which we have agreed to call simple. All the polygonal figures belon to this class of domains. Applying the technique of approximatin a curvilinear domain by polygonal figures we can easily prove th Green's formula is also valid for any domain bounded by a finit number of piecewise smooth curves.

2. Application of Green's Formula to Computing Areas. Green formula implies some useful formulas for computing the area a domain.

Let G be a simple domain with a boundary L and let S be th area of the domain. Consider the line integral

$$\int\limits_L x\,dy$$

Applying Green's formula to the integral we obtain

$$\int\limits_L x\,dy = \iint\limits_G dx\,dy = S$$

Next, we similarly derive the formula

$$S = -\int\limits_L y\,dx$$

Combining these formulas we deduce another formula for computin areas in which the integrations with respect to x and y are involve symmetrically**:

$$S = \frac{1}{2}\int\limits_L x\,dy - y\,dx \qquad (4.47$$

Example. Compute the area of the domain bounded by the astroi

$$x = a\cos^3 t, \quad y = a\sin^3 t$$

* The notion of an improper integral will be discussed in Chapter 9.

** It is apparent that we can receive infinitely many formulas of the forn

$$S = \int\limits_L P\,dx + Q\,dy$$

To this end it is sufficient to take, for P and Q, any two functions satisfyin the condition $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$.

Solution. Applying formula (4.47) we obtain

$$S=\frac{1}{2}\int\limits_{L} x\,dy-y\,dx=\frac{3}{2}a^2\int\limits_{0}^{2\pi}\sin^2 t\cos^2 t\,[\cos^2 t+\sin^2 t]\,dt=$$

$$=\frac{3}{8}a^2\int\limits_{0}^{2\pi}\sin^2 2t\,dt=\frac{3}{8}\pi a^2$$

§ 4. CONDITIONS FOR A LINE INTEGRAL OF THE SECOND TYPE BEING PATH-INDEPENDENT. INTEGRATING TOTAL DIFFERENTIALS

1. Statement of the Problem. When studying the examples of line integrals in § 2 we have noticed that in certain cases a line integral

$$\int\limits_{AB} P\,dx+Q\,dy$$

is independent of the shape of the curve AB and is determined only by the position of its end-points, that is assumes the same values for all the paths of integration joining the fixed points A and B. Here we are going to establish the general conditions guaranteeing such independence of a line integral of the particular choice of the path. This question is closely related to the problem of finding a function of two independent variables from its total differential which we are also going to consider here.

2. The Case of a Simply Connected Domain. We remind the reader (see Chapter 3) that a plane region (domain) G is said to be *simply connected* if each closed contour L lying in the interior of the domain bounds a (finite) part of the plane entirely belonging to G.

Theorem 4.5. Let functions $P(x, y)$ and $Q(x,y)$ and their partial derivatives $\frac{\partial P}{\partial y}$ and $\frac{\partial Q}{\partial x}$ be defined and continuous in a bounded closed simply connected domain G. Then the following four conditions are equivalent to each other (in the sense that the validity of each condition implies the fulfilment of the other three):

1. *The integral*

$$\oint P\,dx+Q\,dy$$

taken over an arbitrary closed contour lying within G is equal to zero.

2. *The integral*

$$\int\limits_{AB} P\,dx+Q\,dy$$

is independent of the particular choice of the path of integration connecting the points A and B (which are considered to be fixed but can be chosen arbitrarily in the domain G).

3. *The expression $P\,dx + Q\,dy$ is the total (exact) differential of a single-valued function defined on the domain G.*

4. *The relation*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \tag{4.48}$$

holds everywhere in the domain G.

Proof. We shall prove the theorem by following the logical scheme

$$1 \to 2 \to 3 \to 4 \to 1$$

i.e. show that the first condition implies the second condition, the second implies the third one, the third implies the fourth and the fourth, in its turn, implies the first condition. Then the equivalence of all the four conditions will be established.

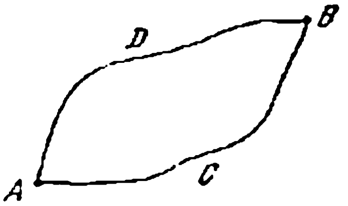


Fig. 4.11

(a) $1 \to 2$. Take two arbitrary paths lying in the domain G and connecting the points A and B, say the paths ACB and ADB shown in Fig. 4.11. Now consider the closed contour $ACBDA$ composed of them. By the hypothesis, the integral taken along any closed contour is equal to zero and thus

$$\int\limits_{ACBDA} P\,dx + Q\,dy = 0$$

But we have

$$\int\limits_{ACBDA} P\,dx + Q\,dy = \int\limits_{ACB} P\,dx + Q\,dy + \int\limits_{BDA} P\,dx + Q\,dy =$$

$$= \int\limits_{ACB} P\,dx + Q\,dy - \int\limits_{ADB} P\,dx + Q\,dy$$

and consequently

$$\int\limits_{ACB} P\,dx + Q\,dy = \int\limits_{ADB} P\,dx + Q\,dy$$

Hence the assertion "$1 \to 2$" has been proved.*

* If the curves ACB and ADB have some points in common other than A and B (see Fig. 4.12) the same result can be established by applying a little more complicated argument.

(b) $2 \rightarrow 3$. Let the integral $\int\limits_{AB} P\,dx + Q\,dy$ be independent of the path of integration. Then, if we fix the point A the integral can be regarded as a single-valued function of the coordinates x and y of the point B:

$$\int\limits_{AB} P\,dx + Q\,dy = U(x, y)$$

Let us show that the function $U(x, y)$ is differentiable and that

$$dU = P\,dx + Q\,dy$$

To prove this it is sufficient to show that the derivatives $\frac{\partial U}{\partial x}$ and $\frac{\partial U}{\partial y}$ exist and are, respectively, equal to $P(x, y)$ and $Q(x, y)$.*

Let us compute the limit

$$\frac{\partial U}{\partial x} = \lim_{\Delta x \to 0} \frac{U(x+\Delta x, y) - U(x, y)}{\Delta x}$$

The quantity $U(x+\Delta x, y) - U(x, y)$ is equal to the integral of $P\,dx + Q\,dy$ taken along an arbitrary path connecting the points

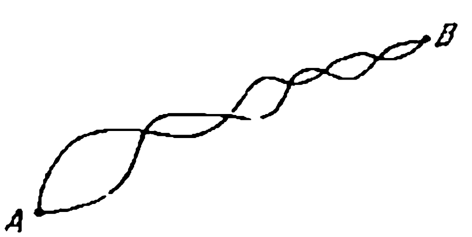


Fig. 4.12

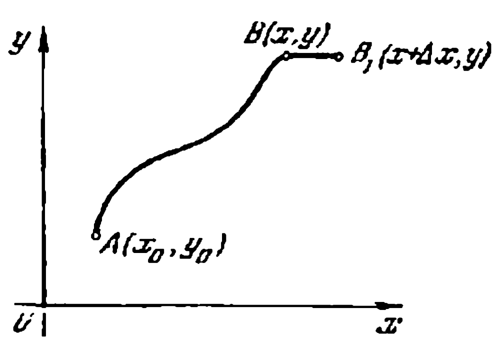


Fig. 4.13

(x, y) and $(x+\Delta x, y)$ since, by the hypothesis, the integral does not depend on the shape of the curve joining the two points. Hence, we can take the path coinciding with the horizontal line segment BB_1 (Fig. 4.13). Therefore, applying the mean value theorem we receive

$$\frac{\Delta U}{\Delta x} = \frac{1}{\Delta x} \int\limits_{BB_1} P\,dx + Q\,dy = \frac{1}{\Delta x} \int\limits_{x,\,y}^{x+\Delta x,\,y} P(x, y)\,dx = P(x+\theta\Delta x, y)$$

where $0 < \theta < 1$. Consequently, we have

$$\frac{\partial U}{\partial x} = \lim_{\Delta x \to 0} P(x+\theta\Delta x, y) = P(x, y)$$

* As is well known, a function possessing continuous partial derivatives is differentiable.

because $P(x, y)$ is continuous. The relation $\frac{\partial U}{\partial y} = Q(x, y)$, is proved similarly.

(c) $3 \to 4$. If the expression $P\,dx + Q\,dy$ is the total differential of a function $U(x, y)$ we have

$$\frac{\partial U}{\partial x} = P, \quad \frac{\partial U}{\partial y} = Q$$

Then the well known theorem asserting that the mixed derivatives $\frac{\partial^2 U}{\partial x\,\partial y}$ and $\frac{\partial^2 U}{\partial y\,\partial x}$ are equal when they are continuous implies

$$\frac{\partial Q}{\partial x} = \frac{\partial^2 U}{\partial x\,\partial y} = \frac{\partial^2 U}{\partial y\,\partial x} = \frac{\partial P}{\partial y}$$

(d) $4 \to 1$. Let the equality $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$ be fulfilled and let L be an arbitrary closed contour lying in the domain G. The domain being (by the hypothesis) simply connected, the part of the plane bounded by the contour L belongs to the domain G in which the functions P and Q and their derivatives are defined and continuous. Therefore, by Green's formula, the line integral

$$\int\limits_L P\,dx + Q\,dy$$

can be transformed into the corresponding double integral

$$\int\limits_L P\,dx + Q\,dy = \iint\limits_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy$$

where D is the domain bounded by the contour L. By virtue of relation (4.48), the integral on the right-hand side is equal to zero. Consequently, we have

$$\int\limits_L P\,dx + Q\,dy = 0$$

for any closed contour L lying within G. The proof of the theorem has thus been completed.

3. Reconstructing a Function from Its Total Differential. In proving Theorem 4.5 we have incidentally solved the following problem (which will be again encountered in § 2, Sec. 4 of Chapter 6): given an expression

$$P\,dx + Q\,dy$$

it is necessary to find a function whose total differential coincides with the expression. In this section we shall limit ourselves to the case when the functions P and Q and their partial derivatives $\frac{\partial P}{\partial y}$

and $\frac{\partial Q}{\partial x}$ are continuous over a simply connected domain G. As has been proved, in these conditions, the expression $P\,dx + Q\,dy$ is the total differential of a function of two arguments if and only if the equality

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

holds (see Theorem 4.5).

Furthermore, in the same proof we have shown that if the above equality is fulfilled the relation

$$dU = P\,dx + Q\,dy \tag{4.49}$$

is satisfied by the function

$$U(x, y) = \int\limits_{(x_0,\,y_0)}^{(x,\,y)} P\,dx + Q\,dy$$

Finally, the formula of finite increments (e.g. see [8], Chapter 8,

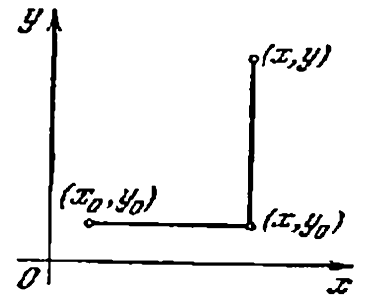


Fig. 4.14

§ 9) suggests that two functions having the same total differential may only differ by a constant term. Consequently, the formula

$$U(x, y) = \int\limits_{(x_0,\,y_0)}^{(x,\,y)} P\,dx + Q\,dy + C \tag{4.50}$$

where (x_0, y_0) is a fixed point and C is an arbitrary constant describes a one-parameter family of functions which contains all the functions satisfying condition (4.49). The integral entering into equality (4.50) being path-independent, we can take at pleasure the curve connecting the points (x_0, y_0) and (x, y). For example, it is convenient to choose, as the path of integration, the broken line composed of the horizontal and the vertical line segments shown in Fig. 4.14.* When the path of integration is chosen in this way equality (4.50) takes the form

$$U(x, y) = \int\limits_{(x_0,\,y_0)}^{(x,\,y_0)} P\,dx + \int\limits_{(x,\,y_0)}^{(x,\,y)} Q\,dy + C$$

* Provided the segments belong to the domain G.

The initial point (x_0, y_0) can be arbitrarily taken within the domain in which the functions P and Q are defined. A change of the position of the point (x_0, y_0) is obviously equivalent to a variation of the value of the arbitrary constant C.

Practically, to determine a function from its total differential it is convenient to apply the following technique. If we have

$$\frac{\partial U}{\partial x} = P, \quad \frac{\partial U}{\partial y} = Q \tag{4.51}$$

then, integrating the first equality with respect to x and considering the variable y entering into it to be a parameter, we obtain

$$U(x, y) = \int P\,dx + f_1 \tag{4.52}$$

where f_1 is independent of x (but, generally speaking, may depend on y, i.e. $f_1 = f_1(y)$). Further, integrating the second equality (4.51) with respect to y while x, in its turn, is regarded as a parameter, we receive

$$U(x, y) = \int Q\,dy + f_2 \tag{4.53}$$

where $f_2 = f_2(x)$. If we now match the functions $f_1(y)$ and $f_2(x)$ in such a way that the right-hand sides of relations (4.52) and (4.53) coincide this will result in a function whose total differential is equal to the expression $P\,dx + Q\,dy$.

Example. Let

$$dU = (2xy + 1)\,dx + (x^2 + 3y^2)\,dy$$

Integrating the coefficient in dx with respect to x we derive

$$\int (2xy + 1)\,dx = x^2y + x + f_1(y) \tag{4.54}$$

The integration of the coefficient in dy with respect to y results in

$$\int (x^2 + 3y^2)\,dy = x^2y + y^3 + f_2(x) \tag{4.55}$$

The right-hand sides of equalities (4.54) and (4.55) coincide if we put

$$f_1(y) = y^3 + C \quad \text{and} \quad f_2(x) = x + C$$

Thus, we see that

$$U = x^2y + x + y^3 + C$$

4. Line Integrals in a Multiply Connected Domain. In the proof of Theorem 4.5 we have taken advantage of the fact that the domain G has been supposed to be simply connected when, on the basis of the condition

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \tag{4.56}$$

we have established the validity of the equality

$$\oint_L P\,dx+Q\,dy=0 \tag{4.57}$$

for any closed contour L lying in G.

Now we shall consider a simple example indicating that, generally speaking, equality (4.57) is not implied by condition (4.56) in the case of a multiply connected domain. Let

$$J=\int_C \frac{-y}{x^2+y^2}\,dx+\frac{x}{x^2+y^2}\,dy \tag{4.58}$$

The integrand does not make sense at the point (0, 0) and therefore we shall delete a neighbourhood of the origin of coordinates. In the plane with the neighbourhood deleted (this plane is now an infinite multiply connected domain) the coefficients in dx and dy are continuous and possess continuous partial derivatives. Besides, we have

$$\frac{\partial}{\partial y}\left(\frac{-y}{x^2+y^2}\right)=\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right)$$

But integral (4.58) taken along a closed contour turns out to be different from zero in the general case. For instance, if C is the circle determined by the equations

$$x=\cos t,\qquad y=\sin t$$

we obtain

$$J=\int_C \frac{-y}{x^2+y^2}\,dx+\frac{x}{x^2+y^2}\,dy=\int_0^{2\pi} dt=2\pi \tag{4.59}$$

Now let us investigate the general properties of an integral

$$\int P\,dx+Q\,dy$$

in case the functions P and Q satisfy the condition*

$$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}$$

but the domain G they are defined in is multiply connected. For definiteness, let us take the domain G depicted in Fig. 4.15, i.e. the one having three "holes", "lacunas", G_1, G_2 and G_3. Let us first consider a closed contour L which does not contain any lacuna inside it. Then we can apply Green's formula to the integral taken over such a contour and thus we see that the integral is equal to zero.

* As before, we suppose that the functions P, Q, $\frac{\partial P}{\partial y}$ and $\frac{\partial Q}{\partial x}$ are continuous in a bounded closed domain G.

Now let L_1 be a contour enveloping one of the lacunas, say the lacuna G_1 (and not containing the other lacunas). Green's formula is no longer applicable to this case and, generally speaking, the integral taken over such a contour is not equal to zero (see the above example).

Let us show that the value of this integral is independent of the particular choice of such a contour containing the lacuna. Let L_1

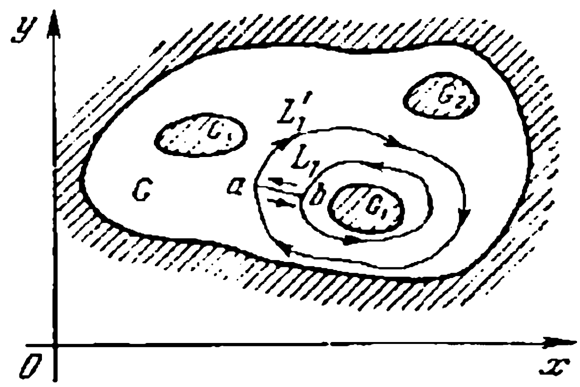


Fig. 4.15

and L_1' be two such contours. Connecting them by an auxiliary curve (ab) we obtain the contour

$$(ab) + L_1 + (ba) - L_1' \tag{4.60}$$

where the minus sign in front of L_1' means that the contour is traversed in the negative direction. Contour (4.60) envelopes no lacunas and hence the integral over it is equal to zero. But the integrals taken along (ab) and (ba) are equal in their absolute values and have the opposite signs. Thus we obtain

$$\int_{L_1} P\,dx + Q\,dy + \int_{-L_1'} P\,dx + Q\,dy = 0$$

that is

$$\int_{L_1} P\,dx + Q\,dy = \int_{L_1'} P\,dx + Q\,dy$$

Consequently, each lacuna G_i $(i = 1, 2, 3)$ in the domain G can be characterized by a certain number ω_i, namely by the value of the line integral $\oint P\,dx + Q\,dy$ taken round an arbitrary closed contour containing the lacuna inside it and not enveloping any other lacuna. Now we can put down the general expression for the integral $\oint P\,dx + Q\,dy$ taken over an arbitrary closed contour L lying in G. Suppose the contour L passes around the first lacuna k_1 times, around the second lacuna k_2 times and around the third one k_3 times. Here, for k_i $(i = 1, 2, 3)$ we each time take into account the direction in which the moving point of the curve L passes around

the ith lacuna, and thus every number k_i is understood as an algebraic sum equal to the difference between the number of times the contour passes around the lacuna in the counterclockwise direction and the number of times it passes around it in the clockwise direction. Then we obviously have

$$\oint_L P\,dx + Q\,dy = k_1\omega_1 + k_2\omega_2 + k_3\omega_3$$

If we make the cuts I, II and III, shown in Fig. 4.16, in the domain G we obtain a simply connected domain in which we can construct the single-valued function

$$U(x, y) = \int_{(x_0,\, y_0)}^{(x,\, y)} P\,dx + Q\,dy \tag{4.61}$$

But, according to what has been said above, the values of the function differ, on the opposite edges of the slit I, by the quantity ω_1,

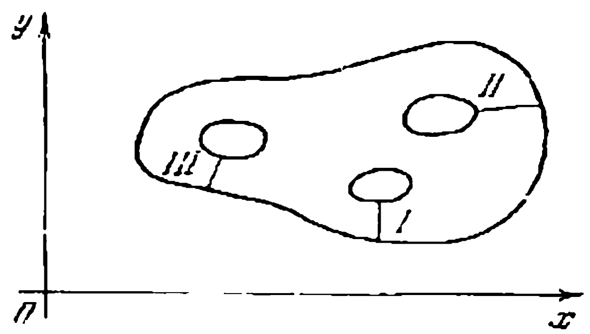


Fig. 4.16

on the edges of the slit II by ω_2 and on the edges of III by ω_3. If we do not make the cuts expression (4.61) again represents a function whose total differential is equal to $P\,dx + Q\,dy$ but in this case the function is multiple-valued. Its values, at a fixed point, corresponding to the contours passing around the lacunas several times differ from each other by terms of the form

$$k_1\omega_1 + k_2\omega_2 + k_3\omega_3$$

where the numbers k_1, k_2 and k_3 can be arbitrary integers (positive, negative or zero).*

It appears clear that all that has been said here is automatically extended to the general case of an arbitrary number of lacunas.

* It may turn out, of course, that all the numbers ω_i are equal to zero in a particular case. Then the function $U(x, y) = \int_{(x_0,\, y_0)}^{(x,\, y)} P\,dx + Q\,dy$ is single valued even when the cuts are not made, and all the assertions of Theorem 4.5 remain true for such a multiply connected domain

5 Surface Integrals

In some physical problems we encounter functions defined on various surfaces. Examples of such functions are the density of a charge distribution over the surface of a conductor, the intensity of illumination of a surface, the velocity of the particles of a fluid passing through a surface and the like. The present chapter is devoted to studying integrals of functions defined on surfaces, the so-called *surface integrals*, and some of their applications.

The theory of surface integrals is in many respects analogous to the theory of line integrals presented in the foregoing chapter. In particular, we shall distinguish between the surface integrals of the first and the second types.

When introducing the definition of a surface integral we shall use some notions concerning surfaces which were discussed in §§ 3 and 4 of Chapter 3 and, particularly, the notion of area of a curvilinear surface.

§ 1. SURFACE INTEGRAL OF THE FIRST TYPE

1. Definition of Surface Integral of a Scalar Function. Let Σ be a piecewise smooth surface bounded by a piecewise smooth contour L.* Consider a bounded function $f(M)$ defined at the points of the surface. Break up the surface Σ into parts $\Sigma_1, \Sigma_2, \ldots, \Sigma_n$ (Fig. 5.1) by means of piecewise smooth curves and denote the areas of the parts as σ_i $(i = 1, 2, \ldots, n)$. Next we choose a point M_i in each part Σ_i and form the sum

$$T = \sum_{i=1}^{n} f(M_i)\,\sigma_i \tag{5.1}$$

which will be referred to as an *integral sum* corresponding to the function $f(M)$ (for the partition $\{\Sigma_i\}$ $(i = 1, 2, \ldots, n)$ of the surface Σ and for the given choice of the points M_i).

---

* In particular, the surface Σ may be closed and have no boundary.

We introduce the following

Definition. *If the integral sums T tend to a finite limit as the maximal of the diameters of the parts Σ_i of the surface Σ tends to zero the limit is called the* ***surface integral of the first type of the function*** *$f(M)$* ***over the surface*** *Σ and is denoted by the symbol*

$$\iint_{\Sigma} f(M)\, d\sigma \tag{5.2}$$

The moving point M on the surface Σ can be determined by its Cartesian coordinates x, y and z. Therefore a function $f(M)$ defined on Σ will also be designated as $f(x, y, z)$ and the corresponding surface integral as $\iint_{\Sigma} f(x, y, z)\, d\sigma$. But when using the last

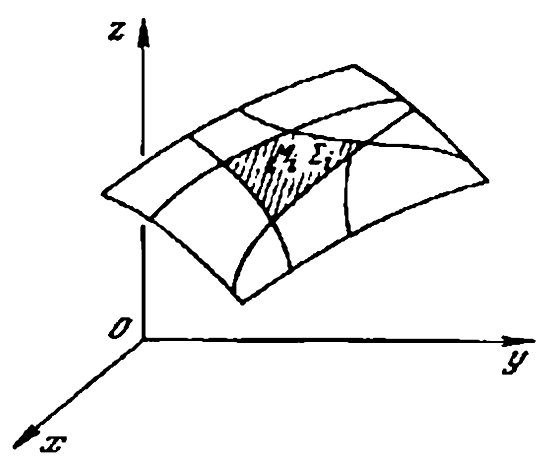


Fig. 5.1

notation one should bear in mind that the variables x, y and z are not independent here but connected by the condition that the point (x, y, z) belongs to the surface Σ.

2. Reducing Surface Integral to Double Integral. We have formulated the definition of the surface integral of the first type and now we are going to discuss the conditions for its existence and the methods of its practical computation.

Both questions are easily answered if we reduce the surface integral to the double integral.

To begin with, we take the simplest case when the surface in question is represented by an equation in Cartesian coordinates.

Theorem 5.1. *Let Σ be a smooth surface determined by an equation $z = z(x, y)$, $(x, y) \in D$, where D is a bounded closed domain, and let $f(x, y, z)$ be a bounded function defined on the surface Σ. Then we have the relation*

$$\iint_{\Sigma} f(x, y, z)\, d\sigma = \iint_{D} f(x, y, z(x, y)) \sqrt{1 + z_x'^2 + z_y'^2}\, dx\, dy \tag{5.3}$$

provided that the integrals entering into it exist. The surface integral on the left-hand side of equality (5.3) *exists if the double integral on the right-hand side does.*

Proof. Divide the surface Σ into n parts Σ_i $(i = 1, 2, \ldots, n)$ by means of piecewise smooth curves. Projecting this partition on

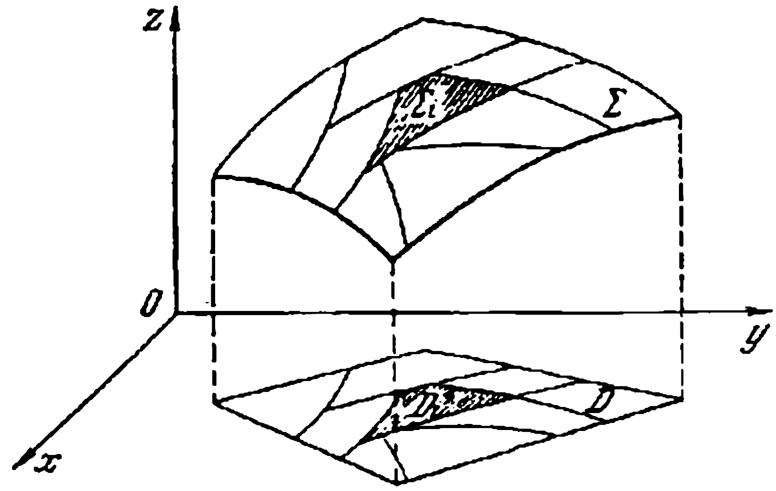


Fig. 5.2

the x, y-plane we obtain a partition of the domain D into squarable parts D_i (see Fig. 5.2) in which the diameter of each element D_i $(i = 1, 2, \ldots, n)$ does not exceed the diameter of the corresponding element Σ_i of the surface Σ.

Consider an integral sum

$$T = \sum_{i=1}^{n} f(x_i, y_i, z_i)\, \sigma_i \tag{5.4}$$

associated with the surface integral $\iint\limits_{\Sigma} f(x, y, z)\, d\sigma$. The area σ_i of the element Σ_i can be written in the form

$$\sigma_i = \iint\limits_{D_i} \sqrt{1 + z_x'^2 + z_y'^2}\, dx\, dy$$

where $z = z(x, y)$. Now taking advantage of the mean value theorem for the double integral of a continuous function* we rewrite the formula for σ_i as

$$\sigma_i = \sqrt{1 + z_x'^2(x_i^*, y_i^*) + z_y'^2(x_i^*, y_i^*)}\, S_i$$

where (x_i^*, y_i^*) is a point belonging to the domain D_i and S_i is the area of the domain. Consequently, integral sum (5.4) can be put down in the form

$$T = \sum_{i=1}^{n} f(x_i, y_i, z(x_i, y_i)) \sqrt{1 + z_x'^2(x_i^*, y_i^*) + z_y'^2(x_i^*, y_i^*)}\, S_i \tag{5.4'}$$

* The surface $z = z(x, y)$ is supposed to be smooth and hence the expression $\sqrt{1 + z_x'^2(x, y) + z_y'^2(x, y)}$ is a continuous function.

Compare this sum with the integral sum

$$\widetilde{T}=\sum_{i=1}^{n} f(x_i, y_i, z(x_i, y_i))\sqrt{1+z_x'^2(x_i, y_i)+z_y'^2(x_i, y_i)}\,S_i \qquad (5.5)$$

corresponding to the double integral on the right-hand side of equality (5.3) (associated with the partition, of the domain D, generated by projecting the partition $\{\Sigma_i\}$ of the surface Σ).

The only distinction between the sums (5.4′) and (5.5) is that in each summand entering into (5.5) the values of the function f and of the expression $\sqrt{1+z_x'^2+z_y'^2}$ are taken at the same point (x_i, y_i) arbitrarily chosen within the element D_i whereas in (5.4′) the values of $\sqrt{1+z_x'^2+z_y'^2}$ are taken at the point (x_i^*, y_i^*) whose position is preassigned by the mean value theorem. Therefore, in the general case, the point (x_i^*, y_i^*) does not coincide with the point (x_i, y_i) although it belongs to the same element D_i.

The function $\sqrt{1+z_x'^2+z_y'^2}$ is continuous in the bounded closed domain D and hence it is uniformly continuous there. Consequently, given any $\varepsilon>0$, there is $\delta_1>0$ such that

$$\left|\sqrt{1+z_x'^2(x_i, y_i)+z_y'^2(x_i, y_i)}-\sqrt{1+z_x'^2(x_i^*, y_i^*)+z_y'^2(x_i^*, y_i^*)}\right|<\varepsilon \qquad (5.6)$$

if the maximal of the diameters of the subdomains D_i is less than δ_1. By the hypothesis, the function $f(x, y, z)$ is bounded, i.e.

$$|f(x, y, z)|\leqslant K=\text{const}$$

and therefore relation (5.6) implies the inequality

$$|T-\widetilde{T}|=\Big|\sum_{i=1}^{n} f(x_i, y_i, z(x_i, y_i))\Big[\sqrt{1+z_x'^2(x_i^*, y_i^*)+z_y'^2(x_i^*, y_i^*)}-$$

$$-\sqrt{1+z_x'^2(x_i, y_i)+z_y'^2(x_i, y_i)}\Big]S_i\Big|\leqslant K\varepsilon\sum_{i=1}^{n} S_i=K\varepsilon S \qquad (5.7)$$

where S is the area of the domain D.

Now we can easily complete the proof of the theorem. If the integral on the right-hand side of (5.3) exists, for every $\varepsilon>0$ there is $\delta_2>0$ such that for any sum $\widetilde{T}$ corresponding to a partition $\{D_i\}$ of the domain D whose elements are of diameters less than δ_2 we have the inequality

$$\left|\iint_D f(x, y, z(x, y))\sqrt{1+z_x'^2(x, y)+z_y'^2(x, y)}\,dx\,dy-\widetilde{T}\right|<\varepsilon \qquad (5.8)$$

Let us take the number $\delta=\min(\delta_1, \delta_2)$ and consider the partitions $\{\Sigma_i\}$ of the surface Σ for which the diameters of all the elements Σ_i are less than δ. Denote by $\{D_i\}$ the partitions of the domain D

corresponding to $\{\Sigma_i\}$. Then the diameter of each D_i is less than δ and consequently inequalities (5.7) and (5.8) are fulfilled. The inequalities imply that

$$\left|\iint_D f(x, y, z(x, y)) \sqrt{1+z_x'^2(x, y)+z_y'^2(x, y)}\, dx\, dy - T\right| < \varepsilon (1+KS)$$

for every partition of the surface Σ whose fineness is small enough. It follows that the limit of the integral sums T exists and equals the integral entering into the right-hand side of relation (5.3), and hence the theorem has been proved.

Corollary. If the surface Σ is smooth and the function $f(x, y, z)$ is continuous the integral

$$\iint_\Sigma f(x, y, z)\, d\sigma$$

is sure to exist.

For, in this case we have a double integral of a continuous function on the right-hand side of equality (5.3) which exists and thus the surface integral on the left-hand side also exists.

Note 1. We have

$$\sqrt{1+z_x'^2+z_y'^2} = \frac{1}{\cos(\mathbf{n}, z)}$$

(see § 3, Sec. 6 in Chapter 3) and therefore equality (5.3) can be rewritten as

$$\iint_\Sigma f(x, y, z)\, d\sigma = \iint_D f(x, y, z(x, y)) \frac{dx\, dy}{\cos(\mathbf{n}, z)} \tag{5.9}$$

If a surface Σ is represented by an equation

$$x = x(y, z)$$

we can interchange the roles of the variables x, y and z and write the relation

$$\iint_\Sigma f(x, y, z)\, d\sigma = \iint_{D_1} f(x(y, z), y, z) \frac{dy\, dz}{\cos(\mathbf{n}, x)} \tag{5.9$_1$}$$

where D_1 is the projection of the surface Σ on the y, z-plane. Similarly, in the case of a surface Σ defined by an equation

$$y = y(z, x)$$

we have the equality

$$\iint_\Sigma f(x, y, z)\, d\sigma = \iint_{D_2} f(x, y(z, x), z) \frac{dz\, dx}{\cos(\mathbf{n}, y)} \tag{5.9$_2$}$$

where D_2 is the projection of Σ on the z, x plane.

Note 2. Suppose a surface Σ is composed of several parts each of which can be represented by an equation of the form

$$x = x(y, z), \quad y = y(z, x) \quad \text{or} \quad z = z(x, y)$$

Then we can take advantage of the fact that the surface integral over Σ is equal to the sum of the integrals taken over the parts the original surface is formed of and apply formulas (5.9), (5.9$_1$) or (5.9$_2$) to each of the integrals separately and thus reduce the integral over Σ to the sum of the double integrals.

In case a surface is represented by parametric equations we can apply arguments essentially the same as above and thus prove the following result.

Theorem 5.1′. Let Σ be a smooth surface represented by a (vector) parametric equation

$$\mathbf{r} = \mathbf{r}(u, v)$$

and $f(x, y, z)$ a bounded function defined on the surface. Then we have the relation

$$\iint_{\Sigma} f(x, y, z)\, d\sigma =$$

$$= \iint_{D} f(x(u, v),\ y(u, v),\ z(u, v)) \sqrt{g_{11}g_{22} - g_{12}^2}\, du\, dv \qquad (5.10)$$

provided that the integrals in (5.10) *exist. The surface integral on the right-hand side exists if the double integral on the left-hand side does.*

Here D is the range of the parameters u, v and g_{11}, g_{12}, g_{22} are the fundamental coefficients of the first order of the surface Σ (see § 4, Sec. 1 in Chapter 3). The expression $\sqrt{g_{11}g_{22} - g_{12}^2}\, du\, dv$ is an element of surface area in the curvilinear coordinates u, v.

Hence, formula (5.10) means that in order to write a surface integral $\iint_{\Sigma} f(x, y, z)\, d\sigma$, taken over a surface Σ which is determined by an equation $\mathbf{r} = \mathbf{r}(u, v)$, in the form of a double integral we must replace the Cartesian coordinates x, y and z of the points of the surface by their expressions in terms of the curvilinear coordinates u and v and substitute the above expression of an element of surface area for $d\sigma$.

Formula (5.3) and formulas (5.9), (5.9$_1$) and (5.9$_2$) are obviously special cases of general formula (5.10). It can be easily shown that these formulas are also valid when the surface is not smooth but piecewise smooth.

3. Some Applications of Surface Integrals to Mechanics. Surface integrals of the first type are frequently encountered in physical problems. For instance, this is the case when we deal with a mass

distribution over a surface and find the coordinates of its centre ‹ gravity, moments of inertia etc. The corresponding formulas aı derived by essentially the same methods as those applied to studyin mass distributions over a plane figure or along a curve (s‹ Secs. 3-5, § 4 in Chapter 1 and Sec. 3, § 1 in Chapter 4) and therefoı we shall only present the final results and leave the computatioı to the reader.

Let a mass of areal density $\rho(x, y, z)$ be distributed over a su face Σ which is smooth or piecewise smooth. We shall suppo: that the function $\rho(x, y, z)$ is continuous on Σ and refer to suc a surface, for brevity, as a *material surface*. Then we have the follo‹ ing results.

(1) The mass μ of the material surface Σ is equal to

$$\mu = \iint_{\Sigma} \rho(x, y, z)\, d\sigma$$

(2) The coordinates of the centre of gravity of the material surfa‹ are expressed by the formulas

$$x_c = \frac{\iint_{\Sigma} x\rho(x, y, z)\, d\sigma}{\iint_{\Sigma} \rho(x, y, z)\, d\sigma}, \quad y_c = \frac{\iint_{\Sigma} y\rho(x, y, z)\, d\sigma}{\iint_{\Sigma} \rho(x, y, z)\, d\sigma}, \quad z_c = \frac{\iint_{\Sigma} z\rho(x, y, z)\, d\sigma}{\iint_{\Sigma} \rho(x, y, z)\, d\sigma}$$

In particular, for a homogeneous surface (with $\rho = \text{const}$) we ha‹

$$x_c = \frac{\iint_{\Sigma} x\, d\sigma}{\iint_{\Sigma} d\sigma}, \quad y_c = \frac{\iint_{\Sigma} y\, d\sigma}{\iint_{\Sigma} d\sigma}, \quad z_c = \frac{\iint_{\Sigma} z\, d\sigma}{\iint_{\Sigma} d\sigma}$$

(3) The moment of inertia of the surface Σ about the z-axis equal to

$$\iint_{\Sigma} (x^2 + y^2)\rho(x, y, z)\, d\sigma$$

and the moments of inertia about the other two axes are express‹ similarly.

4. Surface Integral of a Vector Function. General Concept ‹ Surface Integral of the First Type. We have considered surfa‹ integrals of scalar functions. This notion can be easily generaliz‹ to the vector functions. Let

$$\mathbf{F}(M) = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$$

be a vector function defined on a surface Σ. Let us introduce tl integral of such a function over the surface Σ by putting

$$\iint_{\Sigma} \mathbf{F}(M)\, d\sigma = \mathbf{i}\iint_{\Sigma} P(M)\, d\sigma + \mathbf{j}\iint_{\Sigma} Q(M)\, d\sigma + \mathbf{k}\iint_{\Sigma} R(M)\, d\sigma$$

(5.1

Expression (5.11) will be called the surface integral of the first type of the vector function $\mathbf{F}$ over the surface Σ. The value of such an integral is a vector. The question of existence of a surface integral of the first type of a vector function $\mathbf{F}$, the problem of reducing it to a double integral and the properties of the integral are investigated on the basis of the corresponding facts concerning the integrals of the scalar functions P, Q and R which are the components (coordinates) of the vector $\mathbf{F}$.

To illustrate the application of this notion let us find the force of gravitational attraction with which a material surface attracts a material point.

Let $\rho(x, y, z)$ be the density of mass distribution over a surface Σ and m_0 be a mass concentrated at a point (x_0, y_0, z_0) not belonging

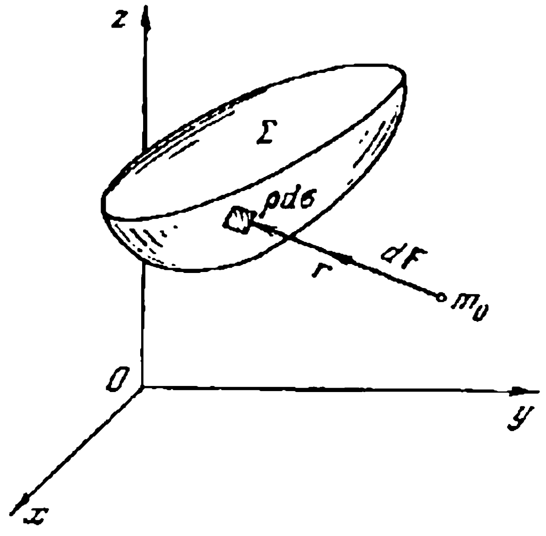


Fig. 5.3

to the surface. An element of area $d\sigma$ carries the elementary mass $\rho(x, y, z)\,d\sigma$ and, by Newton's law of gravitation, the elementary force with which it attracts the mass point m_0 is equal to

$$d\mathbf{F} = \gamma m_0 \rho(x, y, z) \frac{\mathbf{r}}{r^3}\,d\sigma \tag{5.12}$$

Here γ is the constant of gravitation whose numerical value depends on the choice of the system of units and $\mathbf{r}$ is the vector drawn from the point (x_0, y_0, z_0) to the point (x, y, z) (Fig. 5.3). The resultant force $\mathbf{F}$ of attraction of the material point m_0 by the entire surface Σ is equal to the sum of elementary forces (5.12), that is to the surface integral

$$\gamma m_0 \iint_{\Sigma} \rho(x, y, z) \frac{\mathbf{r}}{r^3}\,d\sigma$$

Since we have $\mathbf{r}=(x-x_0)\,\mathbf{i}+(y-y_0)\,\mathbf{j}+(z-z_0)\,\mathbf{k}$, the expression of the force can be written in the form

$$\mathbf{F}=\gamma m_0\Big[\mathbf{i}\iint\limits_{\Sigma}\rho(x,\,y,\,z)\,\frac{x-x_0}{r^3}\,d\sigma+$$

$$+\mathbf{j}\iint\limits_{\Sigma}\rho(x,\,y,\,z)\,\frac{y-y_0}{r^3}\,d\sigma+\mathbf{k}\iint\limits_{\Sigma}\rho(x,\,y,\,z)\,\frac{z-z_0}{r^3}\,d\sigma\Big]$$

The last integral is sure to exist if the surface Σ is smooth or piecewise smooth and the surface density $\rho(x,\,y,\,z)$ is a continuous function on Σ.

An essential feature of the surface integrals of the first type is that each element of integration of the form

$$f(M)\,d\sigma$$

depends solely on the magnitude of the element of area $d\sigma$ and on the value of the function $f(M)$ (which may be scalar or vector) at the corresponding point M but is independent of the orientation of the surface element $d\sigma$ in the surrounding space. This is the case in the physical problems we have considered here because the mass of an element of a material surface at a point M or the force with which the element attracts a material point does not vary if we arbitrarily turn the element about the point M.

But there are problems of another kind in which the orientation of the element $d\sigma$ plays an important role. Such is the problem (to be considered below) of computing the amount of liquid passing through a surface in unit time and some others. These problems lead to another concept of a surface integral, namely to the so-called *surface integral of the second type* which we shall deal with in § 2. As will be shown, there are simple formulas expressing the relationship between the surface integrals of the first and the second types.

§ 2. SURFACE INTEGRAL OF THE SECOND TYPE

1. One-Sided and Two-Sided Surfaces. To give the definition of the surface integral of the second type we must first discuss the question of choosing a side of a surface which is analogous to the problem of introducing an orientation of a curve.

Let Σ be a smooth surface. Take an interior point M_0 on Σ, draw the normal line to the surface at the point and choose one of the two possible directions on the normal. This can be achieved by fixing a certain unit normal vector $\mathbf{n}$ to Σ at the point M_0. Next we consider an arbitrary closed contour C lying on the surface Σ which passes through the point M_0 and has no points in common with the boundary of the surface. Imagine that we are carrying the unit

vector $\mathbf{n}$ along C (starting from the point M_0) in such a way that the vector always remains perpendicular to Σ, i.e. perpendicular to the tangent plane drawn to the surface through the origin of $\mathbf{n}$ (which lies on Σ all the time), and so that the direction of $\mathbf{n}$ continuously varies in this motion. The vector $\mathbf{n}$ always remaining normal to the surface Σ, there are only two possibilities here: (1) after the contour C has been traversed and the moving point (the origin of $\mathbf{n}$) has returned to the point M_0 the new position of the vector $\mathbf{n}$ coincides with the original one; (2) after the contour C has been traversed the vector $\mathbf{n}$ changes its direction to the opposite. Accordingly, we introduce the following

Definition. *A smooth surface Σ is said to be* **two-sided** *if after an arbitrary contour, lying on the surface Σ and having no common points with its boundary, has been described the normal vector to the surface does not change its direction.*

If there is a closed contour such that after it has been traversed the normal vector changes its direction to the opposite the surface is called **one-sided**.

If Σ is a two-sided surface we can choose a unit normal vector $\mathbf{n}(M)$ at each point M of the surface so that $\mathbf{n}(M)$ continuously depends on the point M. To construct such a vector function $\mathbf{n}(M)$ we can choose an initial point M_0 on Σ and one of the two possible directions of the normal vectors $\mathbf{n}(M_0)$ at the point. Then we take an arbitrary point M on Σ, connect it with M_0 by a curve L lying on Σ and carry over the vector $\mathbf{n}$ along L from M_0 to M so that it always remains normal to the surface and its direction continuously varies in this motion on. The vector $\mathbf{n}(M)$ at the point M thus obtained is independent of the choice of the curve L joining the points M_0 and M. For, if two different curves L_1 and L_2 yielded different results we should have formed a closed contour consisting of the curves L_1 and L_2 and thus obtained a closed path C lying on Σ such that after C has been traversed the direction of the normal vector is reversed. But this means that the surface is not two-sided which contradicts the hypothesis.

It clearly follows that, on a two-sided surface Σ, there exist exactly two functions of the type $\mathbf{n}(M)$ continuous throughout Σ. Indeed, each function is completely specified by choosing one of the two possible directions of the normal at an arbitrary initial point on the surface. A function $\mathbf{n}(M)$ of this kind will be referred to as a *continuous field of normals on* Σ. In the case of a one-sided surface it is obviously impossible to construct any continuous field of normals.

When we choose one of the two possible continuous fields of normals on a two-sided surface Σ we thus choose a certain side of the surface (which is seen from the tip of the normal vector).

Examples

1. A plane is a simple example of a two-sided surface. Any part of a plane, e.g. a circle, is also a two-sided surface.

2. Every smooth surface determined by an equation $z = f(x, y)$ is two-sided. In fact, if, at each point of the surface, we take the

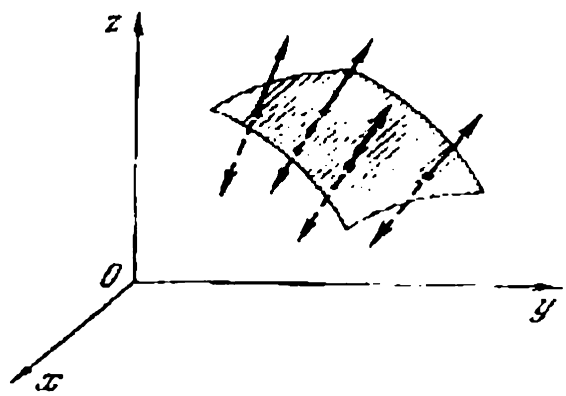


Fig. 5.4

normal vector whose direction forms an acute angle with the positive half-axis z we obtain one (upper) side of the surface, and the other (lower) side if the orientation of the normal vector is reversed (Fig. 5.4).

3. Any closed surface without self-intersections is two-sided, e.g. a sphere, an ellipsoid etc. For instance, if we take, at each point

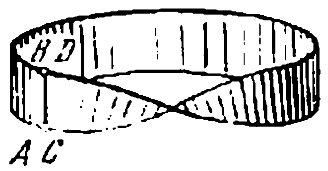


Fig. 5.5

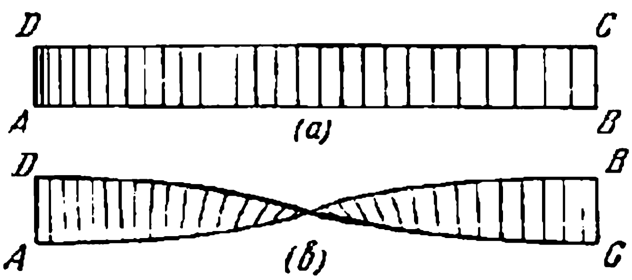


Fig. 5.6

of such a surface, the normal vector directed toward the interior of the solid bounded by the surface we thus indicate the inner side of the surface.

4. The so-called Möbius* strip depicted in Fig. 5.5 is the simplest example of a one-sided surface. It can be obtained by taking a rectangular strip of paper $ABCD$ (Fig. 5.6*a*) and pasting its two ends together after giving it half a twist, i.e. so that the point A coincides with the point C and the point B with the point D (Fig. 5.6*b*). It is easily seen that after we traverse the centre line of the Möbius strip the direction of the normal to the surface is reversed, which means that the surface is in fact one-sided.

---

* Möbius, August Ferdinand (1790-1868), a German mathematician.

Note 1. The two-sided surfaces are also called **orientable** and the process of choosing a certain side of a two-sided surface is referred to as the **orientation** of the surface. The one-sided surfaces are said to be **nonorientable**.

The reader should distinguish between the terms "an orientable surface" (which means that an orientation can be introduced on it, i.e. a certain side of the surface can be chosen) and "an oriented surface" for which a certain side has already been chosen.

Note 2. In contrast to the so-called *local* properties, such as smoothness of a surface, which are determined by the conditions that may

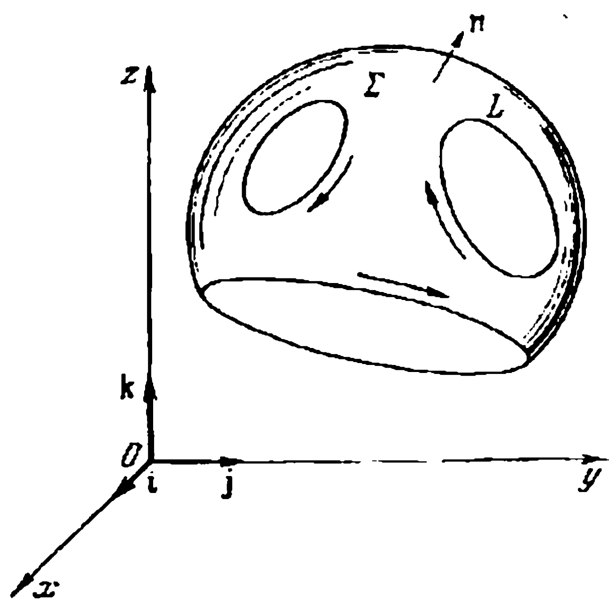


Fig. 5.7

or may not hold at separate points of a surface, the orientability (or nonorientability) of a surface is a *global* property characterizing the surface as a whole. Indeed, a sufficiently small neighbourhood of an interior point of the Möbius strip (or any other surface) proves orientable. In such a neighbourhood it is always possible to construct a continuous field of normals although there exists no such a field on the entire Möbius strip.

The concept of a side of a surface is closely related to the notion of a **coherently (concordantly) oriented** boundary of the surface which we shall need later.*

Let Σ be an oriented surface bounded by a single or several contours. We introduce, for each contour L entering into the boundary, its orientation (coherent with the orientation of the surface Σ) according to the following rule: a direction in which the contour L is described is considered to be **positive** (i.e. coherent with the orientation of Σ) if the surface Σ is always kept on the left of a

* This relationship is dependent on whether the coordinate system taken in the three-dimensional space is right-handed or left handed. In what follows we shall always choose a right-handed coordinate system.

person who is placed on the surface so that the normal vector goes from his feet to his head and who is walking round the contour in this direction (see Fig. 5.7). The opposite direction is referred to as the **negative** one.

If L is an arbitrary closed contour bounding a part of an oriented surface Σ the *positive* direction of traversing the contour, coherent with the orientation of the surface Σ, is again chosen in such a way

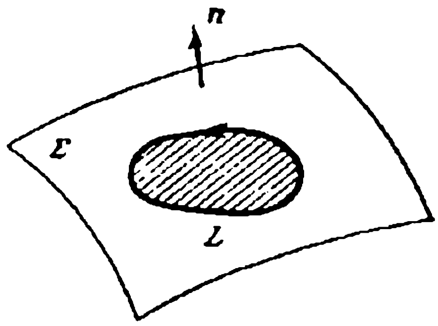


Fig. 5.8

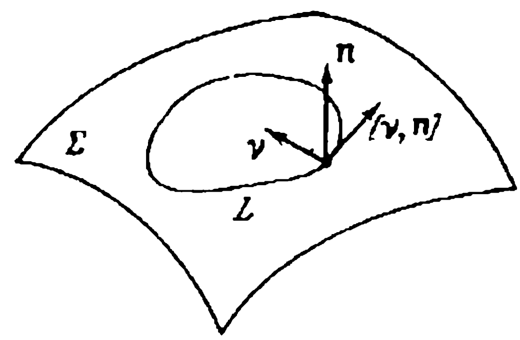


Fig. 5.9

that this part of the surface (it is shaded in Fig. 5.8) remains on the left.* If an oriented plane is taken as Σ the definition of a contour coherently oriented with a surface reduces to the well known rule according to which a contour in the plane is regarded as being positively oriented if it is described in the counterclockwise direction and negatively oriented if otherwise.

Note 3. The rule specifying the coherence between the orientations of a surface Σ and of a contour L entering into the boundary of the surface can also be formulated as follows: let $\mathbf{n}$ be the unit normal vector to the chosen side of the surface Σ at a point M belonging to L, and let $\boldsymbol{\nu}$ be a vector which is perpendicular both to L and to $\mathbf{n}$ and directed toward the surface Σ. Then the positive direction of traversing the contour L coincides with that of the vector $[\boldsymbol{\nu}, \mathbf{n}]$** (see Fig. 5.9).

2. Definition of Surface Integral of the Second Type. Let us first consider a concrete problem involving the notion of a surface integral of the second type, namely the problem of computing the flux of a liquid through a surface.

Let the space (or its part) be filled with a moving liquid (fluid), the velocity of a particle of the liquid passing through an arbitrary

---

* If we take a left-handed coordinate system the rule changes to the opposite, that is the positive direction of tracing a contour L lying on a surface Σ is such that the part of the surface Σ bounded by L remains on the right.

** This rule holds irrespective of whether a right-handed or a left-handed coordinate system has been chosen in the space. The directions of the vectors $\mathbf{n}$ and $\boldsymbol{\nu}$ are independent of the choice of the coordinate system whereas the vector product $[\boldsymbol{\nu}, \mathbf{n}]$ changes its direction to the opposite when a right-handed system is replaced by a left-handed one or vice versa.

point (x, y, z) being specified by a vector $\mathbf{V}(x, y, z)$ with projections (components) $P = P(x, y, z)$, $Q = Q(x, y, z)$ and $R = R(x, y, z)$ on the coordinate axes x, y and z. Let us find the amount Π of liquid passing in unit time through an oriented surface Σ.

Consider an infinitesimal element of area $d\sigma$ of the surface Σ. The quantity of liquid passing through $d\sigma$ in unit time is obviously equal to $d\Pi = V_n\, d\sigma$ where V_n is the projection of the velocity $\mathbf{V}$ on the direction of the normal $\mathbf{n}$ to $d\sigma$ (Fig. 5.10). Expressing V_n as the scalar product of the vector $\mathbf{V}$ by the normal vector $\mathbf{n}$ to $d\sigma$ we obtain

$$d\Pi = [P \cos(\mathbf{n}, x) + Q \cos(\mathbf{n}, y) + R \cos(\mathbf{n}, z)]\, d\sigma \tag{5.13}$$

Formula (5.13) gives an elementary flux of liquid. To obtain

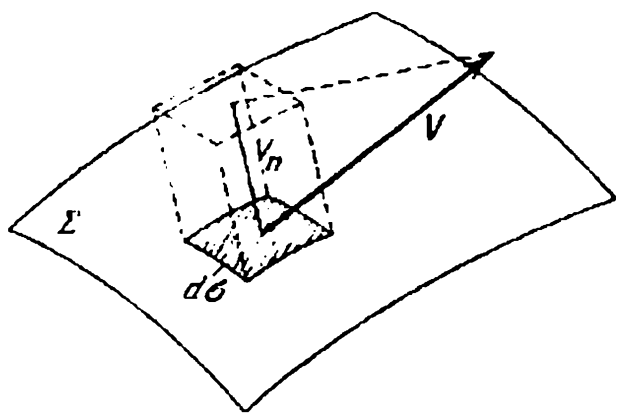


Fig. 5.10

the resultant flux, i.e. the amount of liquid flowing through the whole surface Σ in unit time, we must sum up expressions (5.13) over all the elements $d\sigma$, that is take the integral

$$\Pi = \iint_{\Sigma} [P \cos(\mathbf{n}, x) + Q \cos(\mathbf{n}, y) + R \cos(\mathbf{n}, z)]\, d\sigma$$

According to the definition given in § 1, this is nothing but the surface integral of the first type of the function

$$P \cos(\mathbf{n}, x) + Q \cos(\mathbf{n}, y) + R \cos(\mathbf{n}, z)$$

taken over the surface Σ. But an important thing is that here the integrand depends not only on the vector function (P, Q, R) defined over the surface Σ but also on the direction of the normal at each point of the surface.

Now we proceed to formulate the general definition. Let Σ be a smooth two-sided surface. Fix a certain side of the surface (that is choose one of the two possible fields of normals $\mathbf{n}(M)$) and consider a vector function $\mathbf{A} = (P, Q, R)$ defined on Σ. Let us denote by A_n the projection of the vector $\mathbf{A}$ on the direction of the normal $\mathbf{n}$ to the surface at an arbitrary point (x, y, z). The projection can

be written in the form

$$A_n = P\cos(\mathbf{n}, x) + Q\cos(\mathbf{n}, y) + R\cos(\mathbf{n}, z)$$

where $\cos(\mathbf{n}, x)$, $\cos(\mathbf{n}, y)$ and $\cos(\mathbf{n}, z)$ are the cosines of the angles between the direction of the normal and the directions of the coordinate axes, i.e. the components of the unit normal vector $\mathbf{n}$.

The integral

$$\iint_{\Sigma} [P\cos(\mathbf{n}, x) + Q\cos(\mathbf{n}, y) + R\cos(\mathbf{n}, z)]\, d\sigma \qquad (5.14)$$

will be called the **surface integral of the second type of the vector function** $\mathbf{A} = (P, Q, R)$ **over the surface** Σ (or, strictly speaking over the chosen side of the surface Σ) and will be denoted as

$$\iint_{\Sigma} P\,dy\,dz + Q\,dz\,dx + R\,dx\,dy$$

Thus, by definition, we have the relation

$$\iint_{\Sigma} P\,dy\,dz + Q\,dz\,dx + R\,dx\,dy$$
$$= \iint_{\Sigma} [P\cos(\mathbf{n}, x) + Q\cos(\mathbf{n}, y) + R\cos(\mathbf{n}, z)]\, d\sigma \qquad (5.15)$$

If we pass to the integration over the other side of the surface the components of the unit normal vector change their signs to the opposite and hence we have the same for integral (5.14). It should be noted that for a one-sided surface the notion of a surface integral of the second type is not introduced.

To achieve the generality guaranteeing the possibility of applying the notion of a surface integral of the second type to a wide variety of problems it is advisable to include the integrals over the surfaces having self-intersections (an analogous situation was encountered in the theory of line integrals).

Note 1. If $d\sigma$ is an infinitesimal element of area of a surface Σ the expressions

$$\cos(\mathbf{n}, x)\, d\sigma, \quad \cos(\mathbf{n}, y)\, d\sigma \quad \text{and} \quad \cos(\mathbf{n}, z)\, d\sigma$$

are, respectively, the projections of the element $d\sigma$ on the y, z-, z, x- and x, y-planes (see Fig. 5.11). This is why we denote them as $dy\,dz$, $dz\,dx$ and $dx\,dy$.

Note 2. We have defined the surface integral of the second type on the basis of the notion of the surface integral of the first type. But the surface integral of the second type can also be defined directly by means of the corresponding integral sums, which is performed as follows.

For brevity, let us consider only one of the projections of a continuous vector function (P, Q, R), say R. Take an oriented smooth surface Σ and form a partition of the surface into parts Σ_i. Choosing an arbitrary point (x_i, y_i, z_i) in each part we compose the integral sum

$$\sum_{i=1}^{n} R(x_i, y_i, z_i) S_i \tag{5.16}$$

where S_i is the area of the projection of Σ_i on the x, y-plane taken with the sign + if, at the points belonging to Σ_i, the normal to the surface forms an acute angle with the positive direction of the

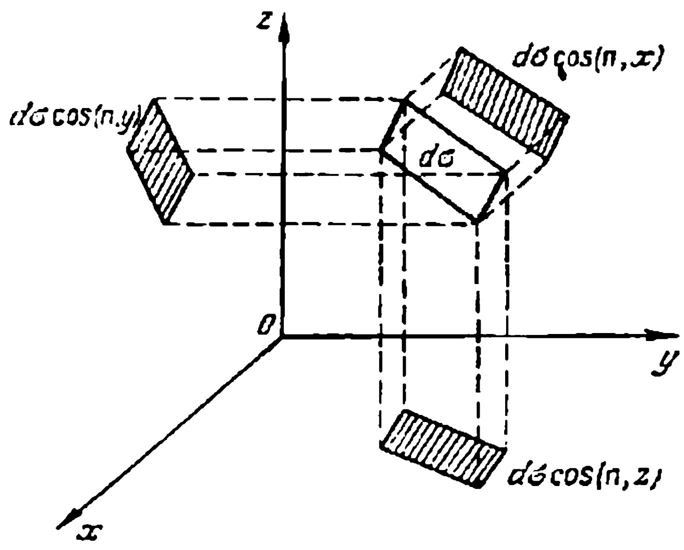


Fig. 5.11

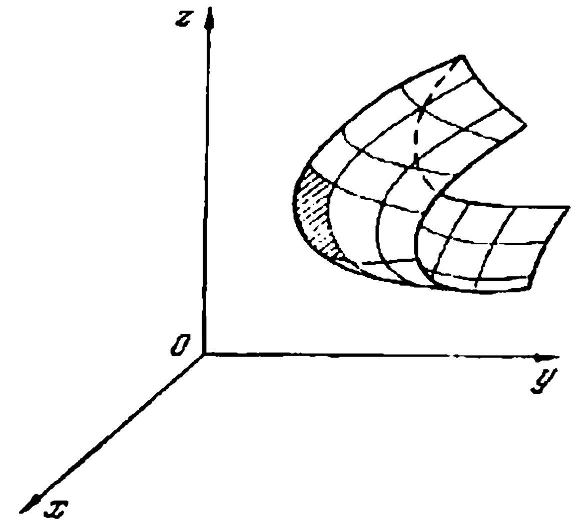


Fig. 5.12

z-axis and the sign — if otherwise, i.e. if the angle is obtuse at each point of the element Σ_i.* We can easily verify that, for a continuous function $R(x, y, z)$ and a smooth surface Σ, the limit of integral sums (5.16) (as the partitions of the surface are infinitely refined) exists and is equal to the integral

$$\iint\limits_{\Sigma} R(x, y, z)\, dx\, dy$$

(compare this with the definition of the line integral of the second type given in § 2, Sec. 2 of Chapter 4).

* A partition of the surface Σ may include "irregular" elements, i.e. such that the angle (n, z) is acute at some of its points and obtuse at the other (see Fig. 5.12). We can either exclude the partitions containing such elements or attribute an arbitrary sign to the areas of their projections because this does not affect the result since the sum of the areas of the projections of these elements is negligibly small.

We can similarly define, by means of the corresponding integral sums, the integrals

$$\iint_{\Sigma} P(x, y, z)\, dy\, dz \quad \text{and} \quad \iint_{\Sigma} Q(x, y, z)\, dz\, dx$$

and, consequently, write the general integral

$$\iint_{\Sigma} P\, dy\, dz + Q\, dz\, dx + R\, dx\, dy$$

as the sum of the integrals of these three forms.

Note 3. The distinction between the surface integrals of the first and the second types is that in the latter each area element $d\sigma$ is in fact regarded not as a scalar quantity but as a vector $d\sigma$ directed along the normal $\mathbf{n}$ to the surface Σ and having the components

$$d\sigma \cos(\mathbf{n}, x), \quad d\sigma \cos(\mathbf{n}, y) \quad \text{and} \quad d\sigma \cos(\mathbf{n}, z)$$

Accordingly, the surface integral of the second type of a vector function $\mathbf{A} = (P, Q, R)$ is often written in the form

$$\iint_{\Sigma} (\mathbf{A}, d\sigma) \tag{5.17}$$

which is equivalent to

$$\iint_{\Sigma} (\mathbf{A}, \mathbf{n})\, d\sigma \tag{5.18}$$

Note 4. Besides integrals (5.18) we encounter, in some problems, integrals of the form

$$\iint_{\Sigma} [\mathbf{A}, \mathbf{n}]\, d\sigma \tag{5.19}$$

The value of such an integral is not a scalar but a vector. The computation of integral (5.19) obviously reduces to the separate integrations of the components (projections) of the vector $[\mathbf{A}, \mathbf{n}]$. Here the integrand also depends, as in integral (5.18), on the normal $\mathbf{n}$ to the surface Σ, and therefore integral (5.19) should be naturally regarded as a surface integral of the second type (but as a "vector-valued" one, in contrast to "scalar" integral (5.18)).

3. Reducing Surface Integral of the Second Type to Double Integral. The definition of a surface integral of the second type and Theorem 5.1 immediately imply the following result.

Let Σ be a smooth (or piecewise smooth) surface determined by an equation

$$z = z(x, y)$$

and let $R(x, y, z)$ be a bounded function defined on Σ. Then, for the surface integral of the second type $\iint\limits_{\Sigma} R(x, y, z)\, dx\, dy$ taken over the upper side of the surface Σ, we have the relation

$$\iint\limits_{\Sigma} R(x, y, z)\, dx\, dy = \iint\limits_{D} R(x, y, z(x, y))\, dx\, dy \qquad (5.20)$$

(where D is the projection of the surface Σ on the x, y-plane) provided that the integrals entering into (5.20) exist. The surface integral on the left-hand side exists if the double integral on the right-hand side of (5.20) exists.

Actually, the surface integral can be written in the form

$$\iint\limits_{\Sigma} R(x, y, z) \cos(\mathbf{n}, z)\, d\sigma$$

Applying formula (5.9) to this expression we obtain the relation we set out to prove.

Thus, in order to reduce a surface integral $\iint\limits_{\Sigma} R(x, y, z)\, dx\, dy$, taken over the upper side of a surface Σ determined by an equation $z = z(x, y)$, to a double integral we must substitute the corresponding function $z = z(x, y)$ for z into the integrand and replace the integration over the surface Σ by the integration over the projection D of Σ on the x, y-plane.

If the integral is taken over the lower side of the surface Σ we obviously have

$$\iint\limits_{\Sigma} R(x, y, z)\, dx\, dy = -\iint\limits_{D} R(x, y, z(x, y))\, dx\, dy$$

We similarly derive the formulas

$$\iint\limits_{\Sigma} P(x, y, z)\, dy\, dz = \pm\iint\limits_{D_1} P(x(y, z), y, z)\, dy\, dz \qquad (5.21)$$

and

$$\iint\limits_{\Sigma} Q(x, y, z)\, dz\, dx = \pm\iint\limits_{D_2} Q(x, y(z, x), z)\, dz\, dx \qquad (5.22)$$

where in the former Σ is understood as a surface represented by an equation $x = x(y, z)$ and in the latter as a surface determined by an equation $y = y(z, x)$. Accordingly, the plus sign is taken if the normal to the surface $x = x(y, z)$ $(y = y(z, x))$ forms an acute angle with the positive direction of the x-axis (y-axis) and the minus sign if the angle is obtuse. The symbols D_1 and D_2 desig-

nate, respectively, the projections of the surface Σ on the y, z- and z, x-axes.

Formula (5.20) can also be applied to reducing a surface integral to a double one when an oriented surface Σ is composed of several pieces each of which is determined by an equation of the form $z = z(x, y)$. In this case the integral should be written as the sum of the integrals corresponding to the pieces and then formula (5.20) should be separately applied to each integral.

Exercise. Rewrite the integral

$$J = \iint_{\Sigma} R(x, y, z)\,dx\,dy$$

taken over the outer side of the sphere

$$x^2 + y^2 + z^2 = a^2$$

in the form of a sum of double integrals.

Answer.

$$J = \iint_{x^2+y^2 \leqslant a^2} R\left(x, y, \sqrt{a^2 - x^2 - y^2}\right) dx\,dy -$$

$$- \iint_{x^2+y^2 \leqslant a^2} R\left(x, y, -\sqrt{a^2 - x^2 - y^2}\right) dx\,dy$$

Here the first summand is equal to the integral taken over the upper side of the upper hemisphere and the second, with the minus sign prefixed to it, is equal to the integral taken over the lower side of the lower hemisphere because the two hemispheres, oriented in this way, constitute the outer side of the entire sphere.

We have shown the way of reducing a surface integral of the second type, taken over a surface determined by an equation in Cartesian coordinates, to a double integral. For a surface represented parametrically, the application of Theorem (5.1′) immediately implies the following result.

If a smooth (or piecewise smooth) surface Σ is represented by a parametric equation

$$\mathbf{r} = \mathbf{r}(u, v)$$

and (P, Q, R) is a bounded vector function defined on Σ we have the relation

$$\iint_{\Sigma} P\,dy\,dz + Q\,dz\,dx + R\,dx\,dy =$$

$$= \iint_{D} [P\cos(\mathbf{n}, x) + Q\cos(\mathbf{n}, y) + R\cos(\mathbf{n}, z)]\sqrt{g_{11}g_{22} - g_{12}^2}\,du\,dv \tag{5.23}$$

where D is the range of the parameters u, v and g_{11}, g_{12}, g_{22} are the fundamental quantities of the first order of the surface Σ, provided that the integrals entering into formula (5.23) exist. The surface integral on the left-hand side of the formula exists if the double integral on the right-hand side does.

Expression (5.23) can be transformed to another form. As is known (see § 3, Sec. 5 in Chapter 3),

$$\cos(\mathbf{n}, x) = \frac{A}{\sqrt{A^2+B^2+C^2}}, \quad \cos(\mathbf{n}, y) = \frac{B}{\sqrt{A^2+B^2+C^2}},$$
$$\cos(\mathbf{n}, z) = \frac{C}{\sqrt{A^2+B^2+C^2}} \qquad (5.24)$$

where

$$A = \begin{vmatrix} \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{vmatrix}, \quad B = \begin{vmatrix} \frac{\partial z}{\partial u} & \frac{\partial x}{\partial u} \\ \frac{\partial z}{\partial v} & \frac{\partial x}{\partial v} \end{vmatrix}, \quad C = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix}$$

and

$$\sqrt{g_{11}g_{22}-g_{12}^2} = \sqrt{A^2+B^2+C^2}$$

Therefore formula (5.23) can be put down as

$$\iint_{\Sigma} P\,dy\,dz + Q\,dz\,dx + R\,dx\,dy = \iint_{D} [PA + QB + RC]\,du\,dv \qquad (5.25)$$

where

$$P = P(x(u, v), y(u, v), z(u, v)),$$
$$Q = Q(x(u, v), y(u, v), z(u, v)) \quad \text{and}$$
$$R = R(x(u, v), y(u, v), z(u, v))$$

Equalities (5.20)-(5.22) are obviously special cases of general formula (5.23).

§ 3. OSTROGRADSKY THEOREM

1. Derivation of Ostrogradsky Theorem. In the foregoing chapter we established a formula connecting a double integral over a plane domain with a line integral taken along its boundary (Green's formula). Here we are going to deduce a similar formula expressing a relationship between a triple integral over a space figure and a surface integral taken over the outer side of the surface bounding the figure. The result we are speaking of is called the *Ostrogradsky theorem (formula)*.*

* Ostrogradsky, Mikhail Vasilyevich (1801-1862), a prominent Russian mathematician. This formula (also called the *divergence theorem*) was published in his article *On Heat Theory* in 1828. It is sometimes called the Gauss theorem although Gauss obtained it considerably later (in 1841).

For convenience, we introduce the following terminology. A spatial domain V bounded by two piecewise smooth surfaces Σ_1 and Σ_2 determined by equations

$$z = z_1(x, y) \quad \text{and} \quad z = z_2(x, y) \quad (z_2(x, y) \geqslant z_1(x, y)) \tag{5.26}$$

and by a lateral cylindrical surface Σ_3 with elements (generators) parallel to the z-axis will be referred to as a domain *regular in the z-direction*. The surfaces $z = z_1(x, y)$ and $z = z_2(x, y)$ will be,

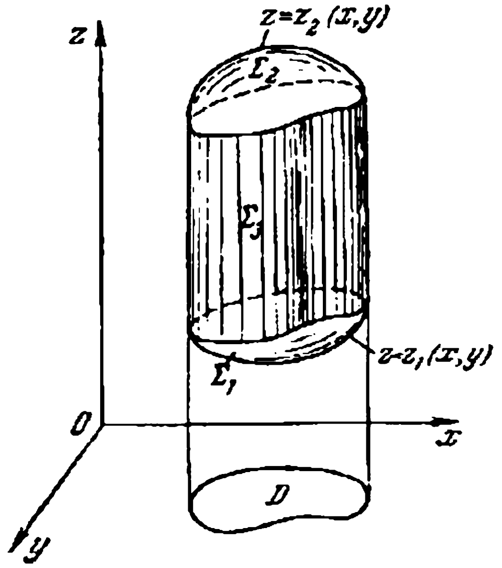


Fig. 5.13

respectively, called the lower and the upper (curvilinear) bases of the domain* (see Fig. 5.13). Similarly, a domain bounded by two piecewise smooth surfaces

$$x = x_1(y, z) \quad \text{and} \quad x = x_2(y, z) \quad (x_2(y, z) \geqslant x_1(y, z))$$

and by a cylindrical surface with generators parallel to the x-axis will be called a domain *regular in the x-direction*. A domain *regular in the y-direction* is defined analogously.

Finally, a domain V will be called *simple* if it is possible to divide it into a finite number of domains regular in the z-direction and also into a finite number of domains regular in the other two directions.

Let now V be a domain, regular in the z-direction, with bases Σ_1 and Σ_2 represented by equations (5.26) and a lateral cylindrical surface Σ_3. The union of the three surfaces Σ_1, Σ_2 and Σ_3 forms the whole boundary of the domain V. We denote the boundary by Σ and consider its outer side. Take a function $R(x, y, z)$ defined

---

* This definition also includes the cases where there can be no lateral surface Σ_3. For instance, a three-dimensional sphere (ball) is considered to be a domain, regular in the z-direction, whose bases are its lower hemisphere Σ_1 and upper hemisphere Σ_2 and whose lateral surface Σ_3 is degenerated into the equator (a ball is also a domain regular in the x-direction and y-direction).

in the domain V (including its boundary) and having the continuous partial derivative $\frac{\partial R}{\partial z}$ in the closure of V. We obviously have the equality

$$\int_{z_1(x,\,y)}^{z_2(x,\,y)} \frac{\partial R}{\partial z}\, dz = R\,(x,\, y,\, z_2\,(x,\, y)) - R\,(x,\, y,\, z_1\,(x,\, y))$$

Let us integrate this equality over the projection D of the domain V on the x, y-plane and replace the threefold iterated integral thus obtained by the corresponding triple integral:

$$\iiint\limits_V \frac{\partial R}{\partial z}\, dx\, dy\, dz = \iint\limits_D R\,(x,\, y,\, z_2\,(x,\, y))\, dx\, dy - \\ - \iint\limits_D R\,(x,\, y,\, z_1\,(x,\, y))\, dx\, dy \qquad (5.27)$$

The first integral on the right-hand side of (5.27) can be written (see formula (5.20)) in the form of the surface integral of the function $R\,(x,\, y,\, z)$ taken over the upper side of the surface

$$z = z_2\,(x,\, y)$$

Similarly, the second integral $\iint\limits_D R\,(x,\, y,\, z_1\,(x,\, y))\, dx\, dy$ can be regarded as the surface integral of the function $R\,(x,\, y,\, z)$ taken over the upper side of the surface $z = z_1\,(x,\, y)$ or as the integral over the lower side of the same surface $z = z_1\,(x,\, y)$ with the minus sign attached to it. Hence we obtain

$$\iiint\limits_V \frac{\partial R}{\partial z}\, dx\, dy\, dz = \iint\limits_{\Sigma_2} R\, dx\, dy + \iint\limits_{\Sigma_1} R\, dx\, dy \qquad (5.28)$$

where the first integral on the right-hand side is taken over the upper side of the surface Σ_2 and the second over the lower side of the surface Σ_1. Adding the integral

$$\iint\limits_{\Sigma_3} R\, dx\, dy$$

(which is obviously equal to zero) taken over the outer side of the lateral surface Σ_3 to the right-hand side of formula (5.28) we obtain, on the right-hand side, a surface integral over the outer side of the entire surface Σ bounding the domain V. Thus, we arrive at the following relation:

$$\iiint\limits_V \frac{\partial R}{\partial z}\, dx\, dy\, dz = \iint\limits_\Sigma R\, dx\, dy = \iint\limits_\Sigma R \cos\,(\mathbf{n},\, z)\, d\sigma \qquad (5.29)$$

Relation (5.29) is also valid for any domain V which can be broken up into a finite number of domains regular in the z-direction. Indeed, let us divide such a domain V into parts V_i regular in the z-direction and write the relation of form (5.29) for each part. Next we sum up all the relations. Then we obtain, on the left-hand side, the triple integral of $\frac{\partial R}{\partial z}$ taken over the whole domain V and, on the right-hand side, the sum of the surface integrals over all the parts of the surface Σ bounding the domain V and over the surfaces breaking up the domain V into the parts V_i, each of the latter integrals being taken twice, that is over one side of the corresponding surface and over its other side. Therefore, after the summation has been performed all the integrals, taken over the surfaces by which the domain V is divided into the parts V_i, mutually cancel out and consequently we derive the formula

$$\iiint\limits_V \frac{\partial R}{\partial z}\, dx\, dy\, dz = \iint\limits_\Sigma R\, dx\, dy \tag{5.30}$$

Let now V be a domain regular in the x-direction, i.e. one bounded by piecewise smooth surfaces (bases)

$$x = x_1(y, z) \quad \text{and} \quad x = x_2(y, z)$$

and by a lateral cylindrical surface with elements parallel to the x-axis. Consider a function $P(x, y, z)$ which is defined and continuous together with its partial derivative $\frac{\partial P}{\partial x}$ in the domain V (including its boundary). Applying arguments similar to those presented above we obtain the equality

$$\iiint\limits_V \frac{\partial P}{\partial x}\, dx\, dy\, dz = \iint\limits_\Sigma P\, dy\, dz \tag{5.31}$$

which also remains valid when the domain V consists of a finite number of domains regular in the x-direction.

We similarly obtain the relation

$$\iiint\limits_V \frac{\partial Q}{\partial y}\, dx\, dy\, dz = \iint\limits_\Sigma Q\, dz\, dx \tag{5.32}$$

for an arbitrary domain V which can be divided into a finite number of parts regular in the y-direction where $Q(x, y, z)$ is a function defined and continuous together with its partial derivative $\frac{\partial Q}{\partial y}$ in the closure of V.

Finally, let us take a simple domain V and consider three functions P, Q and R which are continuous in this domain (including its boundary) and possess the continuous partial derivatives $\frac{\partial P}{\partial x}$,

$\frac{\partial Q}{\partial y}$ and $\frac{\partial R}{\partial z}$ in the closure of V. Then all three relations (5.30), (5.31) and (5.32) are fulfilled. Adding them together we obtain the equality

$$\iiint_V \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right) dx\,dy\,dz=\iint_\Sigma P\,dy\,dz+Q\,dz\,dx+R\,dx\,dy \tag{5.33}$$

or

$$\iiint_V \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right) dx\,dy\,dz=$$

$$=\iiint_\Sigma [P\cos(\mathbf{n},x)+Q\cos(\mathbf{n},y)+R\cos(\mathbf{n},z)]\,d\sigma \tag{5.33'}$$

Formula (5.33) (or (5.33′)) expresses the above mentioned **Ostrogradsky theorem.**

Note. In deriving the Ostrogradsky formula we have supposed that the functions P, Q, R and their partial derivatives $\frac{\partial P}{\partial x}$, $\frac{\partial Q}{\partial y}$, $\frac{\partial R}{\partial z}$ are continuous (and, consequently, bounded) in a closed simple domain. But, applying arguments similar to those presented in connection with Green's formula (see Note 1 in § 3 of Chapter 4) we can prove the validity of the Ostrogradsky formula under more general conditions. Namely, the Ostrogradsky theorem remains true if

1. V is a bounded domain whose boundary consists of a finite number of piecewise smooth surfaces.

2. The functions $P(x, y, z)$, $Q(x, y, z)$ and $R(x, y, z)$ are continuous and, consequently, bounded in the closure of the domain V.

3. The derivatives $\frac{\partial P}{\partial x}$, $\frac{\partial Q}{\partial y}$, $\frac{\partial R}{\partial z}$ exist and are continuous in the interior of the domain V (but do not necessarily satisfy this condition on its boundary) and the integral

$$\iiint_V \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right) dx\,dy\,dz$$

(which is understood as an improper triple integral* in case condition 3 is violated on the boundary) exists.

2. Application of Ostrogradsky Theorem to Evaluating Surface Integrals. Expressing Volume of a Space Figure in the Form of a

---

* Improper multiple integrals will be treated in Chapter 9.

Surface Integral. Formula (5.10) established in § 1 makes it possible to reduce a surface integral of the second type to the corresponding double integral. But there are cases when this way of computing a surface integral proves to be practically inconvenient. In particular, it is sometimes advisable to reduce a surface integral over a closed surface to a triple integral by applying the Ostrogradsky theorem.

Examples

1. Evaluate the integral

$$J=\iint_{\Sigma} x^3\,dy\,dz+y^3\,dz\,dx+z^3\,dx\,dy$$

over the sphere $x^2+y^2+z^2=a^2$.

Solution. Taking advantage of the Ostrogradsky theorem we obtain

$$J=3\iiint_{x^2+y^2+z^2\leqslant a^2}(x^2+y^2+z^2)\,dx\,dy\,dz$$

Now introducing the spherical coordinates we receive

$$J=3\int_0^{2\pi}d\varphi\int_0^{\pi}d\theta\int_0^{a}r^4\sin\theta\,dr=\frac{4}{5}\pi a^5$$

2. Evaluate the integral

$$J=\iint_{\Sigma} z\,dy\,dz+x\,dz\,dx+y\,dx\,dy$$

taken over a closed surface Σ.

Solution. By the Ostrogradsky theorem, the integral reduces to the triple integral (over the domain bounded by the surface Σ) whose integrand is identically equal to zero. Hence, we have $J=0$ for any closed surface Σ.

In the preceding chapter we showed that Green's formula made it possible to express the area of a plane figure as a line integral along its boundary (see formula (4.47)). Similarly, the Ostrogradsky theorem yields an expression of the volume of a space figure V in the form of a surface integral over the closed surface Σ bounding the figure. In fact, let us choose three functions P, Q and R so that

$$\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}=1$$

Then we obtain the relation

$$\iint_{\Sigma} P\,dy\,dz+Q\,dz\,dx+R\,dx\,dy=\iiint_{V}dx\,dy\,dz=V$$

where V is the volume of the domain bounded by Σ. The integral is taken here over the outer side of the surface Σ. In particular, putting

$$P=\frac{1}{3}x, \quad Q=\frac{1}{3}y \quad \text{and} \quad R=\frac{1}{3}z$$

we arrive at a convenient computational formula

$$V=\frac{1}{3}\iint_{\Sigma} x\,dy\,dz+y\,dz\,dx+z\,dx\,dy \tag{5.34}$$

§ 4. STOKES' THEOREM

1. Derivation of Stokes' Formula. *Stokes* formula* expresses a relationship between surface integrals and line integrals. It generalizes Green's theorem, the latter being a special case of the former when the surface in question is a part of the x, y-plane. Like Green's formula and Ostrogradsky's formula, Stokes' formula is widely applied in mathematical analysis and its applications.

Suppose we are given a smooth oriented surface Σ bounded by a coherently oriented contour Λ (see § 2, Sec. 1). Let a vector function (P, Q, R) be defined in a three-dimensional domain in which the surface Σ is strictly contained and let the functions P, Q, R and their first-order partial derivatives be continuous in the domain. We shall transform the line integral

$$\int_{\Lambda} P\,dx+Q\,dy+R\,dz \tag{5.35}$$

taken along the contour Λ into a surface integral over the surface Σ.

We first consider the case when the surface Σ is represented in Cartesian coordinates by an equation

$$z=z(x,\ y)$$

Denote by D the projection of the surface Σ on the x, y-plane, and let L be the boundary of D, that is the projection of the contour Λ (see Fig. 5.14). The transformation of line integral (5.35) into a surface integral will be performed according to the scheme

$$\int_{\Lambda}\to\int_{L}\to\iint_{D}\to\iint_{\Sigma}$$

which means that we shall first transform the line integral over the space curve Λ into a line integral along the plane contour L, then

* Stokes, George Gabriel (1819-1903), an English mathematician and physicist.

reduce it (by means of Green's formula) to a double integral over the domain D and, finally, transform the latter to a surface integral over Σ.

Next we proceed to calculate. To begin with, let us take an integral of the form

$$J_1 = \int_\Lambda P\,dx$$

Observe that we have

$$J_1 = \int_\Lambda P(x, y, z)\,dx = \int_L P(x, y, z(x, y))\,dx$$

because the contour Λ lies on the surface Σ determined by the equation $z = z(x, y)$. Now, applying Green's formula we obtain

$$J_1 = \int_L P(x, y, z(x, y))\,dx = -\iint_D \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial y}\right) dx\,dy \qquad (5.36)$$

where P is a composite function of x and y and therefore its derivative with respect to y has been computed in accordance with the well known rule for differentiating a composite function.

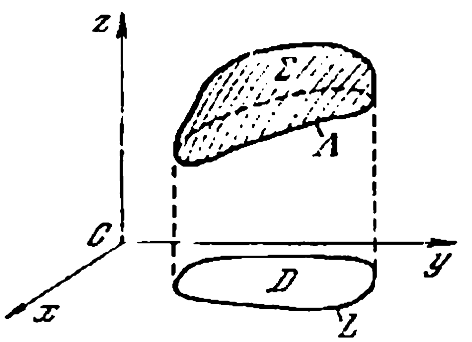


Fig. 5.14

Using expressions (3.36) for the direction cosines of the normal to a surface (represented by an equation $z = z(x, y)$) we find that

$$\frac{\partial z}{\partial y} = -\frac{\cos(\mathbf{n}, y)}{\cos(\mathbf{n}, z)}$$

Thus,

$$J_1 = -\iint_D \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z}\frac{\cos(\mathbf{n}, y)}{\cos(\mathbf{n}, z)}\right) dx\,dy$$

Now, taking advantage of formula (5.20) we can transform the double integral into the corresponding surface integral. This yields

$$J_1 = -\iint_\Sigma \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z}\frac{\cos(\mathbf{n}, y)}{\cos(\mathbf{n}, z)}\right) \cos(\mathbf{n}, z)\,d\sigma =$$

$$= -\iint_\Sigma \left(\frac{\partial P}{\partial y}\cos(\mathbf{n}, z) - \frac{\partial P}{\partial z}\cos(\mathbf{n}, y)\right) d\sigma$$

Hence, we have

$$\int\limits_{\Lambda} P\,dx = \iint\limits_{\Sigma} \left(\frac{\partial P}{\partial z}\cos(\mathbf{n},\,y)-\frac{\partial P}{\partial y}\cos(\mathbf{n},\,z)\right)d\sigma \qquad (5.37)$$

We have supposed here that the surface Σ is represented by an equation of the form $z=z\,(x,\ y)$. But the same result can be obtained for a surface Σ determined by an equation $y=y\,(z,\ x)$. To this end we must consider the projection of Σ on the z, x-plane (instead of x, y-plane) and apply arguments similar to the above. Furthermore, if Σ is a part of a plane perpendicular to the x-axis equality (5.37) remains valid (although in this case Σ cannot be projected either on the x, y-plane or on the z, x-plane so that the correspondence between the points $(x,\ y,\ z)\in\Sigma$ and those of the coordinate planes is one-to-one) since its left-hand and right-hand sides are obviously equal to zero (check it up!). Finally, arguments analogous to those applied to deducing Green's formula and Ostrogradsky's formula show that if the surface Σ is composed of a finite number of parts such that for each of them equality (5.37) holds it also holds for the entire surface Σ. Thus, relation (5.37) has been established for any surface consisting of a finite number of surfaces of the above types. Next we similarly derive the following two equalities analogous to (5.37):

$$\int\limits_{\Lambda} Q\,dy = \iint\limits_{\Sigma} \left(\frac{\partial Q}{\partial x}\cos(\mathbf{n},\,z)-\frac{\partial Q}{\partial z}\cos(\mathbf{n},\,x)\right)d\sigma \qquad (5.38)$$

and

$$\int\limits_{\Lambda} R\,dz = \iint\limits_{\Sigma} \left(\frac{\partial R}{\partial y}\cos(\mathbf{n},\,x)-\frac{\partial R}{\partial x}\cos(\mathbf{n},\,y)\right)d\sigma \qquad (5.39)$$

Adding together equalities (5.37), (5.38) and (5.39) we obtain

$$\int\limits_{\Lambda} P\,dx+Q\,dy+R\,dz=\iint\limits_{\Sigma}\left[\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\cos(\mathbf{n},\,z)+\right.$$

$$\left.+\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\cos(\mathbf{n},\,x)+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\cos(\mathbf{n},\,y)\right]d\sigma \qquad (5.40)$$

which is Stokes' formula that we set out to prove. It can be rewritten in the following form:

$$\int\limits_{\Lambda} P\,dx+Q\,dy+R\,dz=\iint\limits_{\Sigma}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dx\,dy+$$

$$+\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)dy\,dz+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)dz\,dx \qquad (5.41)$$

To remember Stokes' formula one should notice that the first summand under the integral sign on the right-hand side coincides

with the expression under the integral sign in Green's formula and the second and the third summands can be obtained from the first by means of circular permutation of the coordinates x, y, z and the functions P, Q, R.

If the surface Σ is a figure lying in the x, y-plane the integrals involving $dz\,dx$ and $dy\,dz$ vanish and thus Stokes' formula turns into Green's formula.

Note 1. In deriving Stokes' formula we have used a Cartesian coordinate system. But neither the line integral nor the surface integral entering into the formula depends on the way the surface Σ

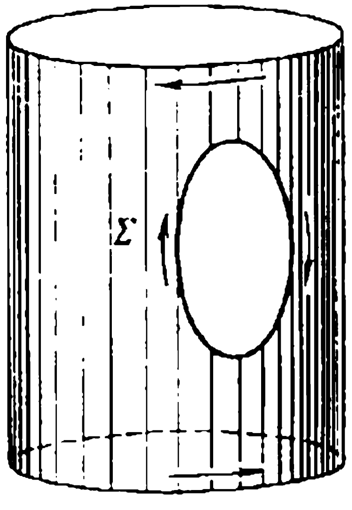


Fig. 5.15

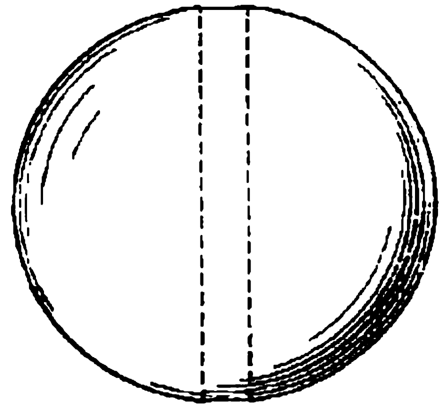


Fig. 5.16

and its boundary Λ are represented. Therefore Stokes' theorem remains true for any other way of representing the surface, including its parametric representation

$$\mathbf{r} = \mathbf{r}\,(u,\ v)$$

Note 2. Stokes' theorem also applies when the boundary Λ of the surface Σ is formed of several separate contours. In this case the integral $\int\limits_{\Lambda} P\,dx + Q\,dy + R\,dz$ should be understood as the sum of the integrals taken over the contours coherently oriented with the surface Σ. For example, if Σ is the lateral surface of a cylinder with an opening (see Fig. 5.15) and if we consider the outer side of the surface, Stokes' theorem expresses the relationship between the integral over Σ and the line integral taken along the three contours, forming the boundary of Σ, in the directions indicated by the arrows in Fig. 5.15.

2. Application of Stokes' Theorem to Investigating Line Integrals in Space. Stokes' theorem has various applications some of which

will be considered in the next chapter. Here we are only going to take advantage of the theorem in order to establish the conditions for a line integral of the second type in space being independent of the path of integration. These conditions generalize the results (obtained by means of Green's formula in § 4 of Chapter 4) concerning the question of path-independence of an integral over a plane curve.

Let us introduce the following

Definition. *A three-dimensional domain V is said to be* ***simply connected*** *if, for any closed contour belonging to V, there exists a surface, with the contour as its boundary, entirely lying in V.*

Examples of simply connected domains are a sphere (ball), the whole space, the domain lying between two concentric spheres etc. As an example of a domain which is not simply connected (such domains are referred to as **multiply connected**) we can take a ball with a cylindrical tunnel passing through it (see Fig. 5.16).

Next we proceed to establish the following result analogous to Theorem 4.5.

Theorem 5.2. *If $P(x, y, z)$, $Q(x, y, z)$ and $R(x, y, z)$ are continuous functions, defined in a bounded closed simply connected domain V, which possess the continuous first-order partial derivatives in the domain the following four assertions are equivalent to each other.*

1. *The integral $\oint P\,dx + Q\,dy + R\,dz$ taken over any closed contour lying inside V is equal to zero.*

2. *The integral $\int\limits_{AB} P\,dx + Q\,dy + R\,dz$ is independent of the path of integration connecting two arbitrary fixed points A and B.*

3. *The expression $P\,dx + Q\,dy + R\,dz$ is the total differential of a single-valued function defined in V.*

4. *The conditions*

$$\frac{\partial Q}{\partial x}=\frac{\partial P}{\partial y}, \quad \frac{\partial R}{\partial y}=\frac{\partial Q}{\partial z}, \quad \frac{\partial P}{\partial z}=\frac{\partial R}{\partial x} \tag{5.42}$$

are fulfilled at each point (x, y, z) of the domain V.

Proof. The theorem is proved according to the scheme $1 \to 2 \to 3 \to 4 \to 1$ which we followed when proving Theorem 4.5. We leave the proof to the reader with the only hint that to deduce condition 1 from condition 4 one must take an arbitrary closed contour Λ lying within V and consider a surface Σ entirely lying in V whose boundary is Λ, such a surface existing because of the condition that V is a simply connected domain. Then the application of Stokes' theorem to the line integral taken over Λ shows that condition (5.42)

implies the relation

$$\int\limits_{\Lambda} P\,dx + Q\,dy + R\,dz = 0$$

If the expression $P\,dx + Q\,dy + R\,dz$ is the total differential of a function $U(x, y, z)$ we can easily derive the formula

$$U(x, y, z) = \int\limits_{(x_0, y_0, z_0)}^{(x, y, z)} P\,dx + Q\,dy + R\,dz + C \qquad (5.43)$$

analogous to formula (4.50) established in § 4 of Chapter 4 for the case of two independent variables. The symbol $\int\limits_{(x_0, y_0, z_0)}^{(x, y, z)}$ designates here the integral along an arbitrary path entirely lying in the domain V and connecting an arbitrary fixed point (x_0, y_0, z_0) with a variable point (x, y, z), C being an arbitrary constant.

If the functions P, Q and R satisfy conditions (5.42) but the domain they are defined in is not simply connected the properties of the integral

$$\int\limits_{AB} P\,dx + Q\,dy + R\,dz$$

resemble those of the line integral $\int\limits_{AB} P\,dx + Q\,dy$ in a plane multiply connected domain. In particular, in the case of a multiply connected domain, expression (5.43) is also a function whose total differential coincides with $P\,dx + Q\,dy + R\,dz$ but in the general case the function may be multiple-valued.

6 Field Theory

The concept of *field* forms the basis for various notions of modern physics. In this chapter we shall present the elements of mathematical theory applied to investigating physical fields.

In physical problems we usually deal with quantities of two basic types, namely scalars and vectors.* Accordingly, we shall consider two types of field, i.e. *scalar fields* and *vector fields*.

§ 1. SCALAR FIELD

1. Definition and Examples of Scalar Field. Let Ω be a domain in space. If there is a correspondence which attributes a number $U(M)$ to each point M of the domain we say that there is a **scalar field** *defined in the domain* Ω.

Examples of scalar fields are a temperature field inside a body subjected to heating (in this case at each point M of the body the corresponding temperature $U(M)$ is specified), field of illumination produced by a light source etc.

The density field of a mass distribution we have already dealt with is an important example of a scalar field. Let us come back to this notion. Suppose that a spatial domain Ω carries a continuously distributed mass. Associating with every subdomain V belonging to Ω the mass contained in V, we arrive at an additive set function $\mu(V)$. If, at each point $M \in \Omega$, the set function possesses the derivative $\frac{d\mu}{dV}$ with respect to volume the function $\rho(M) = \frac{d\mu}{dV}$ is called the *density of mass* and the values of the derivative form a scalar field referred to as the **density field** *of mass distribution*. Similarly, a continuous distribution of an electric charge yields a scalar field of *charge density*. There are many other examples of this kind.

* By the way, only in studying some simpler problems of physics can we limit ourselves to scalar and vector quantities. In many branches of modern theoretical physics, such as electrodynamics, the theory of relativity, the theory of elementary particles etc., an essential role is played by the quantities of a more complicated nature. In our course we shall deal with one type of such quantities, namely with the so-called *tensors* which will be studied in the next chapter.

Besides the fields defined in spatial domains, we often encounter plane scalar fields. An example of such a field is the illumination of a part of a plane produced by a light source.

2. Level Surfaces and Level Lines. Let $U(M)$ be a scalar field. If we introduce a Cartesian coordinate system x, y, z in the domain of definition of the field this field can be represented by a scalar function $U(x, y, z)$ of the coordinates of the moving point M.* In what follows, a function $U(x, y, z)$ of this type will be considered to be continuous and to have continuous partial derivatives of the first order with respect to the variables x, y and z.

The specification of a scalar field by means of a fixed coordinate system and the corresponding function $U(x, y, z)$ is sometimes insufficient for visualizing the structure of the field. To get a more complete description it is convenient to use the so-called level surfaces. A **level surface** *of a scalar field $U(M)$ is a locus of the points at which the field $U(M)$ assumes a given fixed value C.* The equation of a level surface is of the form

$$U(x, y, z) = C^{**} \tag{6.1}$$

It is clear that the level surfaces (corresponding to all the possible values of C) fill the entire domain in which the field is defined and that two surfaces

$$U(x, y, z) = C_1 \quad \text{and} \quad U(x, y, z) = C_2$$

have no points in common for $C_1 \neq C_2$. The specification of all the level surfaces and the corresponding values of C marked on them is equivalent to the specification of the field $U(M)$. The disposition of the level surfaces in space enables us to visualize the structure of the field.

This method of representing a field is especially convenient when the field in question is defined in a plane region. In this case the field is represented by a function $U(x, y)$ of two variables. Generally speaking, an equality of the form $U(x, y) = C$ determines

---

* The character of the function $U(x, y, z)$ is of course dependent not only on the field in question but also on the choice of the coordinate system. If the coordinate system is regarded as being fixed the notion of a spatial scalar field coincides with that of a function of three variables. But nevertheless, to stress that our discussion concerns, as a rule, the quantities which have a certain physical significance independent of the choice of the coordinate system we shall always use the term "field".

** Under the above assumptions concerning the functions of type $U(M)$ equation (6.1) in fact determines a smooth surface provided that, for the given C, there exist points satisfying the equation and the derivatives $\frac{\partial U}{\partial x}$, $\frac{\partial U}{\partial y}$ and $\frac{\partial U}{\partial z}$ do not simultaneously vanish at these points (e.g. see [8], Chapter 14, § 4).

a curve. Such curves are called **level lines** *of the plane scalar field* $U(M)$. Level lines are widely applied in cartography for representing the relief of a terrain. For this purpose, to indicate altitude on a topographic map, the *contour lines* (the *horizontals*) connecting the points of the same elevation are drawn (see Fig. 6.1). This method is also used for representing, on special maps, the distribution of

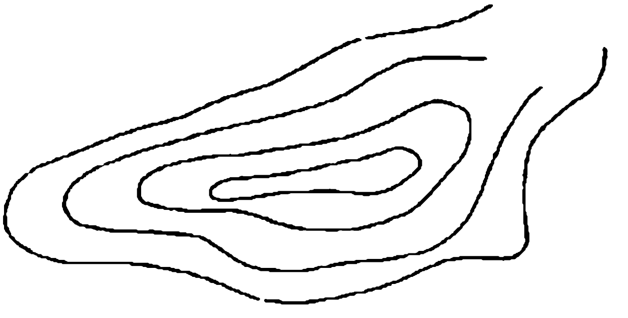


Fig. 6.1

temperature, pressure, amount of precipitation and the like. The corresponding level lines are then referred to as *isoterms* in the case of temperature, *isobars* in the case of pressure etc.

3. Various Types of Symmetry of Field. In many physical problems we deal with fields possessing various types of symmetry. The symmetry properties usually simplify the investigation of such fields. Let us indicate some important special cases.

(a) *Two-dimensional field.* If there is a Cartesian coordinate system in which a scalar field $U(M)$ can be represented by a function dependent not on three but on two coordinates (e.g. by a function $U(x, y)$) the field is said to be **two-dimensional (plane-parallel).** In other words, a scalar field $U(M)$ is called plane-parallel if there is a direction in space such that the field goes into itself when being translated along this direction. The level surfaces of such a field form a family of cylindrical surfaces (Fig. 6.2). In an appropriately chosen coordinate system the family is represented by an equation of the form $U(x, y) = C$.

(b) *Axially symmetric field.* If, for a given field $U(M)$, there exists a cylindrical coordinate system in which the field is represented by a function depending solely on the variables $r = \sqrt{x^2 + y^2}$ and z (but not on the angle φ) the field is said to be **axially symmetric.** This means that a field $U(M)$ is axially symmetric if and only if it goes into itself when being rotated (through an arbitrary angle) about a fixed straight line which is the axis of symmetry of the field. The level surfaces of such a field are obviously surfaces of revolution (Fig. 6.3). In case these surfaces of revolution are circular cylinders (see Fig. 6.4), that is if the field $U(M)$ is represented, in a specifically chosen coordinate system, by a function dependent on only one coordinate r (which is the distance from the point M to the axis of symmetry of the field), $U(M)$ is called a **cylindrical field.**

(c) *Spherical field*. If the values of $U(M)$ depend only on the distance of the point M from a fixed point M_0 the field is said to be spherical. The level surfaces of this field constitute a family of concentric spheres (Fig. 6.5).

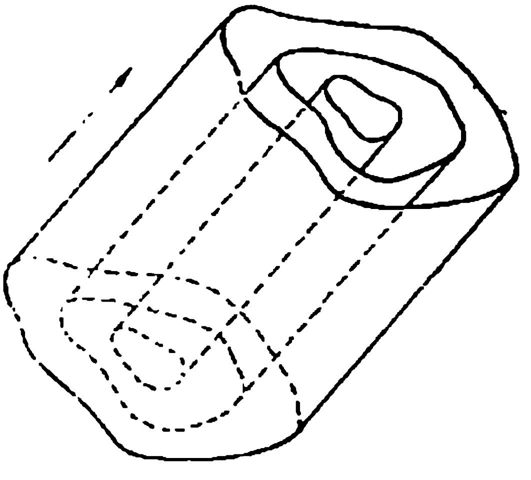


Fig. 6.2

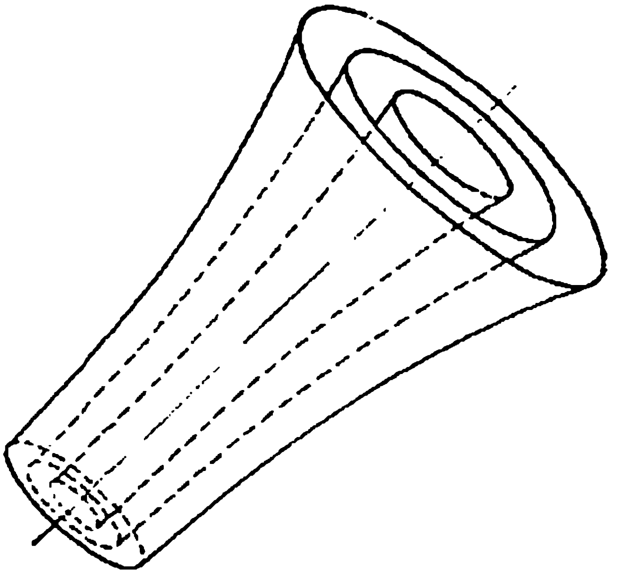


Fig. 6.3

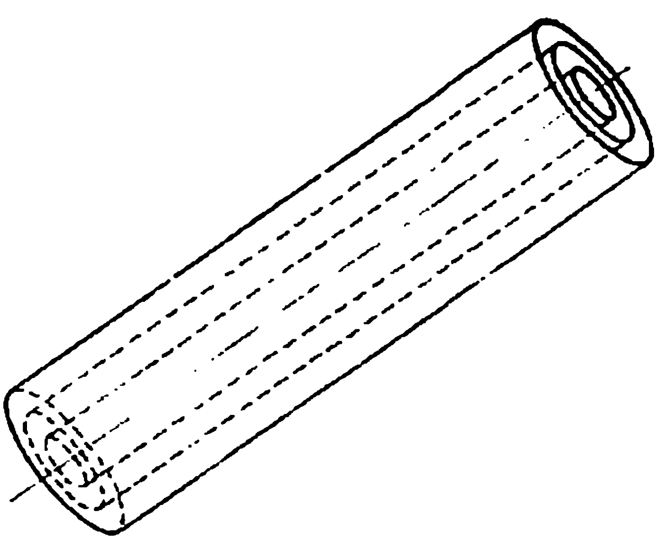


Fig. 6.4

4. Directional Derivative. The application of mathematical analysis to investigating a scalar field $U(M)$ makes it possible to describe its local properties, i.e. the variation of $U(M)$ in passing from a given point M to the points M' lying close to M.

To this end we shall use the notion of the *directional derivative of a field*. Let $U(M)$ be a scalar field. Consider two points M and M' placed close to each other and form the ratio

$$\frac{U(M')-U(M)}{h} \tag{6.2}$$

where h is the length of the line segment MM'. Let the point M' approach M so that the direction of the line segment MM' all the time coincides with the direction of a fixed unit vector $\boldsymbol{\lambda}$. If, in this process, ratio (6.2) tends to a finite limit we call this limit the

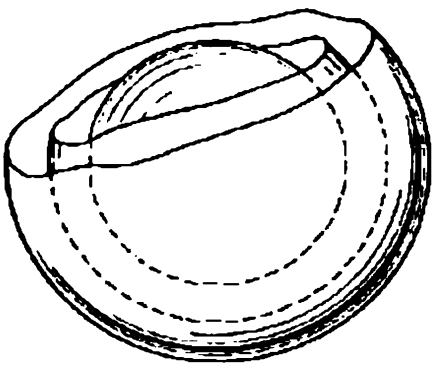


Fig. 6.5

derivative of the scalar field $U(M)$ at the point M along the direction of the vector λ (the directional derivative) and designate it as

$$\frac{\partial U(M)}{\partial \lambda}$$

The derivative $\frac{\partial U}{\partial \lambda}$ characterizes the rate of change of the quantity $U(M)$ in the direction λ.

To compute $\frac{\partial U}{\partial \lambda}$ we choose a coordinate system and represent $U(M)$ as a function $U(x, y, z)$.

Suppose the direction $\boldsymbol{\lambda}$ (i.e. the vector $\boldsymbol{\lambda}$) forms angles α, β and γ with the coordinate axes. Then we have

$$\overrightarrow{MM'} = h(\mathbf{i}\cos\alpha + \mathbf{j}\cos\beta + \mathbf{k}\cos\gamma)$$

and

$$U(M') = U(x + h\cos\alpha,\ y + h\cos\beta,\ z + h\cos\gamma) \qquad (6.3)$$

Hence the derivative $\frac{\partial U}{\partial \lambda}$ coincides with the derivative of composite function (6.3) with respect to h for $h \to 0$. Differentiating we thus obtain

$$\frac{\partial U(M)}{\partial \lambda} = \frac{\partial U(M')}{\partial h}\bigg|_{h=0} = \frac{\partial U}{\partial x}\cos\alpha + \frac{\partial U}{\partial y}\cos\beta + \frac{\partial U}{\partial z}\cos\gamma \qquad (6.4)$$

5. Gradient of Scalar Field. Expression (6.4) can be regarded as a scalar product of two vectors, namely of the unit vector

$$\boldsymbol{\lambda} = (\cos\alpha,\ \cos\beta,\ \cos\gamma)$$

determining the direction in which the derivative $\frac{\partial U}{\partial \lambda}$ is taken and the vector having the components (projections on the x-, y- and z-axes)

$$\frac{\partial U}{\partial x}, \quad \frac{\partial U}{\partial y} \quad \text{and} \quad \frac{\partial U}{\partial z}$$

The latter vector is known as the gradient of the scalar field U and is denoted by the symbol

$$\operatorname{grad} U$$

Thus, we can write

$$\operatorname{grad} U = \left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z} \right) \tag{6.5}$$

and consequently

$$\frac{\partial U}{\partial \lambda} = (\operatorname{gard} U, \lambda) \tag{6.6}$$

Fig. 6.6 visually interprets the expression of the directional derivative in the form of the projection of grad U on the vector λ.

Formula (6.6) can also be written as

$$\frac{\partial U}{d\lambda} = |\operatorname{gard} U| \cos \varphi$$

where φ is the angle between grad U and the unit vector λ. It follows that at every point where grad $U \neq 0$ there is a single direction

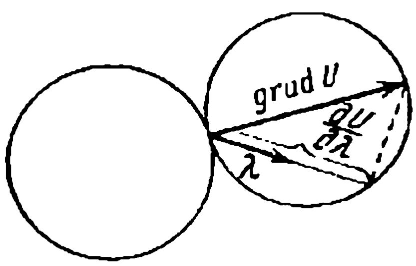


Fig. 6.6

for which $\frac{\partial U}{\partial \lambda}$ assumes its maximum value, i.e. a single direction in which the function U increases with the maximum rate This direction coincides with the direction of the vector grad U. Indeed, for this direction $\varphi = 0$ and consequently

$$\frac{\partial U}{\partial \lambda} = |\operatorname{grad} U|$$

whereas for all the other directions we have

$$\frac{\partial U}{\partial \lambda} = |\operatorname{grad} U| \cos \varphi < |\operatorname{grad} U|$$

Thus we see that the direction of the vector grad U is the one in which the quantity U increases with the maximum rate and that the magnitude (length) of the vector grad U equals the rate of increase of the quantity U in this direction.

But, obviously, neither the direction of maximum rate of increase of a function nor the value of its directional derivative in this direction depends on the choice of the coordinate system. Hence we come to the conclusion that the gradient of a scalar field is specified by the field itself and not by the choice of the coordinate system (although

this assertion is not directly implied by equality (6.5) which we consider to be a definition of the gradient).

As is known, the derivatives $\frac{\partial U}{\partial x}$, $\frac{\partial U}{\partial y}$ and $\frac{\partial U}{\partial z}$ at a point M are the components (projections) of a vector normal to the surface $U(x, y, z) = \text{const}$ passing through the point.* Therefore, the gradient of a field U is directed along the normal to the level surface at each point of the domain where the field is defined.

Every curve for which the tangent line at each point M goes along the vector grad U at the point M will be referred to as a **gradient line of the field** U.** We can now say that the gradient lines of a field U are the curves along which the rate of change of the field is maximal.

It can be proved that if $U(x, y, z)$ is a function possessing continuous partial derivatives up to the second order inclusive for every point M of the domain of definition of the field U there exists a single gradient line passing through the point M. A gradient line is orthogonal, at its every point, to the level surface passing through the point.

§ 2. VECTOR FIELD

1. Definition and Examples of Vector Field. We say that there is a vector field defined in a domain Ω if a certain vector $\mathbf{A}(M)$ is associated with each point M of the domain.

An important example of a vector field which will be many times discussed in what follows is the **field of velocities** of a stationary flow of a liquid. Let a domain Ω be occupied by a liquid (fluid) flowing, at each point, with a velocity $\mathbf{v}$ independent of time (but varying, in the general case, from point to point). If we associate, with each point M of the domain Ω, the vector $\mathbf{v} = \mathbf{v}(M)$ we thus arrive at a vector field which is a field of velocities.

The field of gravitation of a mass distribution is another important example of a vector field. Suppose a mass is distributed in space. Then a material point with unit mass placed at a point M is subjected to a gravitational force. These forces (defined at every point) form a vector field which is called the **field of gravitation** corresponding to the given mass distribution.

---

* Actually, if a straight line λ lies in the tangent plane to the surface $U(x, y, z) = \text{const}$ at the point M the derivative of U in the direction of λ is obviously equal to zero:

$$\frac{\partial U}{\partial \lambda} = (\text{grad } U, \lambda) = 0$$

Thus, the vector grad U is perpendicular to any vector lying in the tangent plane and hence it goes along the normal to the surface.

** Compare this with the general definition of a *vector line* given in § 2.

If there are electric charges distributed in space they act upon unit charge placed at a point M with a certain force $\mathbf{F}(M)$. These forces constitute a vector field which is referred to as an **electrostatic field**.

A field of gravitation and an electrostatic field are examples of a **field of force**.

Let $\mathbf{A}(M)$ be a vector field in space. Taking a Cartesian coordinate system in space we can represent $\mathbf{A}(M)$ as a collection of three scalar functions which are the components of the vector. As a rule, in what follows these components (projections of $\mathbf{A}(M)$ on the coordinate axes) will be denoted by $P(x, y, z)$, $Q(x, y, z)$ and $R(x, y, z)$. In this chapter we shall consider the vector fields whose components are continuous and possess continuous first-order partial derivatives.*

2. Vector Lines and Vector Surfaces. Let a vector field $\mathbf{A}(M)$ be defined in a domain Ω. A curve L lying in Ω is called a **vector line** if the direction of the tangent to the curve at each point coincides with the direction of the vector $\mathbf{A}$ at the point. In particular, if $\mathbf{A}$ is the field of velocities of a stationary flow of a liquid its vector lines are the trajectories of the particles of the liquid.

In some questions related to investigating vector fields the problem of finding a vector line, of a field $\mathbf{A}$, passing through a given point M_0 plays an important role. The problem can be formulated analytically: it is necessary to determine a vector function $\mathbf{r}(t)$ satisfying the conditions

$$\mathbf{r}'(t) = \lambda\mathbf{A} \tag{6.7}$$
$$\mathbf{r}(t_0) = \mathbf{r}_0$$

where $\mathbf{r}_0$ is the radius vector of the initial point M_0, t_0 is the initial instant of time and λ is an arbitrary scalar parameter. It can be shown that if the components P, Q and R of the vector $\mathbf{A}$ are continuously differentiable functions of the coordinates which do not simultaneously vanish, conditions (6.7) in fact determine a single vector line passing through the point M_0.**

A bounded surface Σ lying in a part of space where a vector field $\mathbf{A}$ is defined is said to be a **vector surface** if at each point of the surface Σ the normal to Σ is orthogonal to the vector $\mathbf{A}$ at the point. Thus we can say that every vector surface Σ is made up of vector lines; each vector line either entirely lies on Σ or has no points in common with it.

---

* It is clear that if this condition is fulfilled for one coordinate system it automatically holds for any other coordinate system.

** This is a consequence of the existence and uniqueness theorem for a solution of a system of differential equations with given initial conditions (e.g. see [5], Chapter 1, § 6).

A part of space lying in the domain of definition of a vector field **A** and bounded by a tubular vector surface is referred to as a **tube of the vector field.** Such a tube is entirely composed of vector lines, and every vector line of the field **A** either entirely lies within the tube or is placed outside it.

If a field **A** is thought of as a field of velocities of a stationary fluid flow, a tube of the field is the part of space which is traced by a fixed volume of fluid in the process of motion.

3. Types of Symmetry of Vector Field. As in the case of a scalar field, the investigation of a vector field is simplified when it possesses symmetry properties. Let us enumerate some of the most important special cases.

(a) *Plane-parallel field.* If, for a given vector field **A**, it is possible to choose a Cartesian coordinate system in which the components of the field **A** have the form $P(x, y)$, $Q(x, y)$ and $R(x, y)$, that is are independent of z, the field **A** is said to be **plane-parallel.** If, in addition, we have $R(x, y) = 0$ the field **A** is called **plane.** An example of such a field is the velocity field of a stationary flow of a liquid whose particles have the velocities parallel to a fixed plane and independent of the distance from a particle to the plane (a **plane flow**). The vector lines of such a field are plane curves whose shape and disposition are the same in each plane parallel to the given plane.

(b) *Axially symmetric field.* A vector field **A** is called **axially symmetric** if there exists a cylindrical coordinate system r, φ, z such that at each point M the vector $\mathbf{A}(M)$ depends only on r and z and does not depend on φ. In other words, such a field goes into itself when it is rotated about the z-axis. In case the vector $\mathbf{A}(M)$ depends solely on r the field is said to be **cylindrical**.

(c) *One-dimensional field.* A **one-dimensional field** is characterized by the possibility of choosing a Cartesian coordinate system in which the components of the field have the form $P(x)$, 0, 0. The family of the vector lines of a one-dimensional field obviously coincides with the totality of all straight lines parallel to the x-axis.

4. Field of Gradients. Potential Field. Consider a scalar field $U(M)$. Constructing the vector grad U at every point M we arrive at a vector field which is the **field of gradients** *of the scalar quantity* U. Next we introduce the following

Definition. A vector field $\mathbf{A}(M)$ *is said to be* ***potential*** *if it can be represented as a field of gradients of a scalar field* $U(M)$, *i.e.*

$$\mathbf{A} = \operatorname{grad} U$$

In this case the scalar field U is called the **potential (potential function)** of the vector field **A**.

Let us take the following example. Put $U = f(r)$ where $r = \sqrt{x^2 + y^2 + z^2}$. We thus obtain the spherical field U. To find the gradient of U we differentiate with respect to x and thus obtain

$$\frac{\partial U}{\partial x} = f'(r)\frac{\partial r}{\partial x} = f'(r)\frac{x}{r}$$

and, similarly,

$$\frac{\partial U}{\partial y} = f'(r)\frac{y}{r}, \quad \frac{\partial U}{\partial z} = f'(r)\frac{z}{r}$$

Consequently, we have

$$\operatorname{grad} U = f'(r)\frac{\mathbf{r}}{r}, \qquad \mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \tag{6.8}$$

If a vector field **A** has a potential function this function is uniquely specified, to within an arbitrary constant addend, by the field. Indeed, if scalar fields U and V have the same gradient we can write

$$\operatorname{grad}(U - V) \equiv 0$$

But then the directional derivative of $U - V$ is equal to zero in all directions at each point which immediately implies that

$$U - V = \text{const}$$

The vector lines of a potential field **A** are obviously the gradient lines of its potential U, i.e. the curves along which the rate of change of U is maximal.

It is now natural to ask what are the conditions for a vector field being potential. The answer to the question was in fact found in Chapter 5. Indeed, as was shown (see Theorem 5.2), an expression

$$P\,dx + Q\,dy + R\,dz$$

(where P, Q and R are continuous functions possessing continuous partial derivatives of the first order) is the total differential of a single-valued function $U(x, y, z)$ if and only if P, Q and R satisfy the conditions*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \quad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z} \tag{6.9}$$

But the relation

$$P\,dx + Q\,dy + R\,dz = dU$$

is equivalent to

$$P = \frac{\partial U}{\partial x}, \quad Q = \frac{\partial U}{\partial y}, \quad R = \frac{\partial U}{\partial z}$$

* Here we of course suppose that the domain in which the vector field **A** is considered is simply connected.

and hence conditions (6.9) imply that the field (P, Q, R) is potential.

Thus, for a vector field $\mathbf{A} = (P, Q, R)$ with continuous and continuously differentiable components $P(x, y, z)$, $Q(x, y, z)$ and $R(x, y, z)$ to be potential, it is necessary and sufficient that equalities (6.9) be fulfilled.

If we are given a potential vector field $\mathbf{A}$ the potential function can practically be found from its total differential as was illustrated in the problem considered in § 4 of Chapter 5 (formula (5.43)) for the case of three independent variables and in § 4 of Chapter 4 (formula (4.50)) for two variables.

The notion of a potential field will be discussed again in Sec. 5 of § 4.

Example. Let a mass m be concentrated at the origin O. If we now place unit mass at a point $M(x, y, z)$ it will be acted upon by the gravitational force

$$\mathbf{F} = -\gamma \frac{m}{r^3}\mathbf{r} \quad (\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k})$$

These forces are determined at all points in space and thus form a vector field which is the **field of gravitation of the mass point** m. This field can be represented as the gradient of the function

$$\frac{\gamma m}{r}$$

known as the **Newtonian potential** of the mass point m. To verify that this function is in fact the potential of the field we take advantage of formula (6.8) and thus obtain

$$\operatorname{grad} \frac{\gamma m}{r} = -\gamma \frac{m}{r^3}\mathbf{r}$$

§ 3. FLUX OF VECTOR FIELD. DIVERGENCE

1. Flux of Vector Field Across a Surface. In the foregoing chapter (§ 2) we showed that the amount of fluid passing through a given (oriented) surface Σ in unit time is equal to the integral

$$\iint_{\Sigma} A_n \, d\sigma$$

where A_n is the projection of the velocity vector $\mathbf{A} = (P, Q, R)$ on the (outer) normal to the surface.

We called this quantity the *flux of the liquid* (through the surface Σ).

Now let $\mathbf{A}$ be an arbitrary vector field and Σ be an oriented surface. The surface integral

$$\iint_{\Sigma} A_n \, d\sigma$$

will be called the flux of the vector field $\mathbf{A}$ across the surface Σ.

Thus, if $\mathbf{A}$ is the velocity of a fluid flow the flux of the vector $\mathbf{A}$ across a surface is equal to the quantity of the liquid passing through the surface in unit time. For a vector field of some other nature the flux of the field may have another physical meaning.

Example. Let $U = U(x, y, z)$ be a temperature field inside a physical body and let $\mathbf{A} = \operatorname{grad} U$. Denoting by k the *coefficient of thermal conductivity* we can apply the *Fourier* law of heat propagation* and express the quantity of heat dQ passing in unit time through an element $d\sigma$ of an oriented surface Σ in the form

$$dQ = -k \frac{\partial U}{\partial n} d\sigma \tag{6.10}$$

where $\frac{\partial U}{\partial n}$ is the derivative of the field of temperature in the direction of the (outer) normal to $d\sigma$. (The minus sign on the right-hand side of equality (6.10) is due to the well known fact that heat travels, within a heat conductor, from the regions of higher temperature to those of lower temperature, i.e. in the direction of decrease of U.) Since we have

$$\frac{\partial U}{\partial n} = (\operatorname{grad} U)_n$$

equality (6.10) can be rewritten as

$$dQ = -k (\operatorname{grad} U)_n \, d\sigma$$

whence it follows that the amount of heat passing in unit time across the whole surface Σ is equal to

$$Q = -\iint_{\Sigma} k (\operatorname{grad} U)_n \, d\sigma \tag{6.11}$$

Introducing the vector

$$\mathbf{q} = -k \operatorname{grad} U$$

known as the heat flux vector we obtain

$$Q = \iint_{\Sigma} q_n \, d\sigma$$

* Fourier, Jean Baptiste Joseph (1768-1830), a prominent French mathematician and physicist.

Consequently, the quantity of heat passing through Σ in unit time is equal to the flux of the vector $\mathbf{q}$ across the surface Σ.

2. Divergence. Let $\mathbf{A}$ be a vector field which will be thought of as a velocity field of an incompressible fluid. The liquid being incompressible, the flux

$$\Pi = \iint_{\Sigma} A_n \, d\sigma$$

of the vector $\mathbf{A}$ across a closed surface Σ * is obviously equal to the amount of liquid which is introduced or removed, within the domain Ω (bounded by the surface Σ) in unit time. This quantity characterizes the total capacity of the sources, in case $\Pi > 0$, or the **sinks** (also termed **negative sources**), if $\Pi < 0$, lying in the domain Ω. Consider the ratio

$$\frac{\iint_{\Sigma} A_n \, d\sigma}{V(\Omega)}$$

of the fluid flux through the surface Σ to the volume $V(\Omega)$ of the domain Ω bounded by the surface. It expresses the mean source (sink) density, that is the amount of the liquid introduced (removed) within unit volume of the domain Ω.

Finally, let us consider the limit

$$\lim_{\Omega \to M} \frac{\iint_{\Sigma} A_n \, d\sigma}{V(\Omega)}$$

of the ratio where the symbol $\lim\limits_{\Omega \to M}$ indicates the passage to the limit as the domain Ω is contracted to a fixed point M. This limit is the measure of source (sink) density of the fluid at the point M. It is a scalar quantity which is an important characteristic of the field.

Now we can pass to general definitions.

Let $\mathbf{A}$ be an arbitrary vector field. With every spatial domain Ω bounded by a smooth or piecewise smooth surface Σ (lying in the part of space where the field is defined), we now associate the quantity

$$\iint_{\Sigma} A_n \, d\sigma$$

which is the flux of the vector $\mathbf{A}$ across the outer side of the surface Σ. We thus arrive at a set function $\Phi(\Omega) = \iint_{\Sigma} A_n \, d\sigma$. It can be easily verified that the set function is additive.

---

* Here we consider the outer side of the surface.

Definition. *The derivative of the function* $\Phi(\Omega)$ *with respect to volume, that is the limit*

$$\lim_{\Omega\to M}\frac{\iint\limits_{\Sigma} A_n\,d\sigma}{V(\Omega)} \tag{6.12}$$

is termed the ***divergence of the vector field*** $\mathbf{A}$ *and denoted by the symbol*

$$\operatorname{div}\mathbf{A}$$

Thus, the source density of the field of velocities of a fluid flow which has been considered above is equal to the divergence of the field.

Theorem 6.1. *If* $\mathbf{A}=(P,\ Q,\ R)$ *is a vector field, defined in a domain* Ω, *such that the functions* P, Q, R *are continuous and possess continuous first-order partial derivatives in* Ω, *the divergence* $\operatorname{div}\mathbf{A}$ *exists at all the points of the domain and is expressed, in every Cartesian coordinate system, by the formula*

$$\operatorname{div}\mathbf{A}=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z} \tag{6.13}$$

Proof. Let us apply the Ostrogradsky formula

$$\iint\limits_{\Sigma} A_n\,d\sigma=\iiint\limits_{\Omega}\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)dv$$

By virtue of the theorem on differentiating a triple integral with respect to volume (see § 1, Sec. 5 of Chapter 2), the derivative with respect to volume of the right-hand side of the formula exists and is equal to $\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}$. Consequently, the derivative with respect to volume of the left-hand side also exists and is equal to the same expression. The latter derivative being, by definition, the divergence $\operatorname{div}\mathbf{A}$ of the vector field $\mathbf{A}$, the theorem has thus been proved.

Note. The relation

$$\operatorname{div}\mathbf{A}=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}$$

is often taken as a definition of the divergence. But such definition is less convenient than the one given here because it is based on the choice of the coordinate system and therefore the fact that the sum $\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}$ is independent of the choice of the coordinate system should be additionally proved whereas the independence of expression (6.12) of the choice is quite apparent.

Thus, with every vector field $\mathbf{A}$ whose components are continuous and have continuous partial derivatives of the first order, we

associate the scalar function div A, the divergence of the vector field A. Using this notion we can rewrite the Ostrogradsky formula as follows:

$$\iint_{\Sigma} A_n \, d\sigma = \iiint_{\Omega} \operatorname{div} \mathbf{A} \, dv \tag{6.14}$$

Hence the flux of a vector A through the outer side of every closed surface Σ is equal to the integral of the divergence of the field A taken over the domain bounded by the surface Σ.

3. Physical Meaning of Divergence for Various Types of Field. Examples.

(a) As was shown, for the velocity field A of an incompressible liquid moving in a spatial domain Ω, the expression

$$\iiint_{\Omega} \operatorname{div} \mathbf{A} \, dv$$

is the measure of the total capacity of sources (sinks) placed in the domain Ω, and div A is equal to the source intensity per unit volume. In particular, if A is a field of velocities of an incompressible fluid whose flow has neither sources nor sinks, we have

$$\operatorname{div} \mathbf{A} = 0$$

(b) Let us now consider the field of gravitation of a mass distribution and elucidate the physical significance of the divergence of such a field. To begin with, we take the field produced by a mass m_0 concentrated at a point (x_0, y_0, z_0). Then the force acting upon unit mass placed at a point (x, y, z) is equal to

$$\mathbf{F} = \left(\gamma m_0 \frac{x-x_0}{r^3}, \quad \gamma m_0 \frac{y-y_0}{r^3}, \quad \gamma m_0 \frac{z-z_0}{r^3}\right) \tag{6.15}$$

where $r = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$ and γ is the constant of gravitation whose numerical value depends on the choice of the system of units. In what follows we assume that the system of units is so chosen that $\gamma = 1$.

Let us compute the divergence of force field (6.15). At each point distinct from (x_0, y_0, z_0) we have

$$\frac{\partial}{\partial x}\left(m_0 \frac{x-x_0}{r^3}\right) = m_0 \frac{r^3 - 3(x-x_0)^2 r}{r^6} = m_0 \frac{r^2 - 3(x-x_0)^2}{r^5}$$

and, similarly,

$$\frac{\partial}{\partial y}\left(m_0 \frac{y-y_0}{r^3}\right) = m_0 \frac{r^2 - 3(y-y_0)^2}{r^5}$$

and

$$\frac{\partial}{\partial z}\left(m_0 \frac{z-z_0}{r^3}\right) = m_0 \frac{r^2 - 3(z-z_0)^2}{r^5}$$

Adding together the results we obtain

$$\operatorname{div}\mathbf{F} = m_0 \frac{3r^2 - 3(x - x_0)^2 - 3(y - y_0)^2 - 3(z - z_0)^2}{r^5} = 0$$

But these calculations do not apply in the case of the point (x_0, y_0, z_0) at which the finite value of the divergence cannot be defined. Therefore the value of the integral

$$\iiint\limits_{\Omega} \operatorname{div}\mathbf{F}\, dv$$

(which is an improper integral in this case, see Chapter 9) cannot be obtained by a direct integration if the domain Ω contains the point (x_0, y_0, z_0). Therefore the expression on the right-hand side of Ostrogradsky formula (6.14) is undetermined in this case. But we can easily find the quantity on the left-hand side of the formula, that is the flux of the vector $\mathbf{F}$ across the surface Σ bounding the domain Ω (and attribute the numerical value thus obtained to the integral $\iiint\limits_{\Omega} \operatorname{div}\mathbf{F}\, dv$ when the point (x_0, y_0, z_0) is contained in Ω).

We now proceed to calculate the flux. Let us first take, as the surface Σ, a sphere of radius a with centre at the point (x_0, y_0, z_0). The direction of vector (6.15) coincides with the direction of the normal to the sphere at each point of the sphere. Therefore, in this case, the projection of vector (6.15) on the normal is equal to the length of the vector, i.e. to the constant quantity $\frac{m_0}{a^2}$. Consequently we have

$$\iint\limits_{\Sigma} F_n\, d\sigma = \frac{m_0}{a^2} 4\pi a^2 = 4\pi m_0$$

Substituting any other closed surface Σ_1, containing the point (x_0, y_0, z_0) in the interior of the domain bounded by it, for the sphere Σ we arrive at the same result. In fact, we can choose the sphere Σ with a sufficiently small radius $a > 0$ so that it is entirely contained within Σ_1. Then we have

$$\iint\limits_{\Sigma_1} F_n\, d\sigma - \iint\limits_{\Sigma} F_n\, d\sigma = 0$$

since the left-hand side of this equality is equal to the flux of the vector $\mathbf{F}$ across the boundary of the spatial domain in which

$$\operatorname{div}\mathbf{F} \equiv 0$$

Hence, we obtain

$$\iint\limits_{\Sigma_1} F_n\, d\sigma \qquad \iint\limits_{\Sigma} F_n\, d\sigma$$

Now we shall consider the field of gravitation produced by several mass points. This field is equal to the sum of the fields corresponding to each separate mass. The flux of a sum of fields through a surface being obviously equal to the sum of the fluxes of the summands, it follows that the flux of the field of gravitation, produced by a system of mass points, across a closed surface is equal to the sum of the masses contained inside the surface multiplied by 4π.

Applying the well known technique of passing to the limit from a system of material points to a mass continuously distributed in space with a volume density $\rho(x, y, z)$ we can show* that for a continuous mass distribution the flux of the field of gravitation through a closed surface Σ is also equal to the total mass, contained within this surface, multiplied by 4π. But the total mass can be represented as the integral of the density $\rho(x, y, z)$ taken over the volume Ω bounded by the surface Σ. Therefore, denoting, as before, the vectorial value of the gravitational field at an arbitrary point (x, y, z) by the symbol $\mathbf{F}(x, y, z)$ we can write

$$\iint_{\Sigma} F_n(x, y, z)\, d\sigma = 4\pi \iiint_{\Omega} \rho(x, y, z)\, dv$$

whence

$$4\pi\rho(x, y, z) = \lim_{\Omega \to (x, y, z)} \frac{\iint_{\Sigma} F_n\, d\sigma}{V(\Omega)}$$

The integral on the right-hand side is the divergence of the vector field $\mathbf{F}$. Thus, we finally conclude that the divergence of the field of gravitation produced by a continuous mass distribution is equal to the volume density $\rho(x, y, z)$ of the distribution multiplied by 4π.

(c) The argument which has been applied to the field of gravitation can also be used for investigating the electrostatic field. This results in the Gauss theorem which is widely applied to various problems concerning electrostatic fields, e.g. to the problem of determining the electric field intensity in the capacitors of various types. The theorem states that the divergence of an electrostatic field is equal to the charge density multiplied by 4π.

4. Solenoidal Field. A vector field whose divergence is identically equal to zero is said to be *solenoidal***. As was seen, the velocity

---

* The rigorous justification of such a passage to the limit is dealt with in the so-called potential theory and is based on the theory of integrals involving parameters. The fundamentals of the theory of integrals dependent on a parameter are presented in Chapter 10.

** From the Greek word σωλην tube.

field of an incompressible fluid is an example of a solenoidal field in case there are no sources and sinks, i.e. when there are no points at which the fluid is introduced or removed.

For the solenoidal fields, we have the so-called law of conservation of intensity of a vector tube which we are going to deduce here. Let **A** be a solenoidal field. Consider a tube of the vector field and take its part contained between two sections Σ_1 and Σ_2 (see Fig. 6.7).

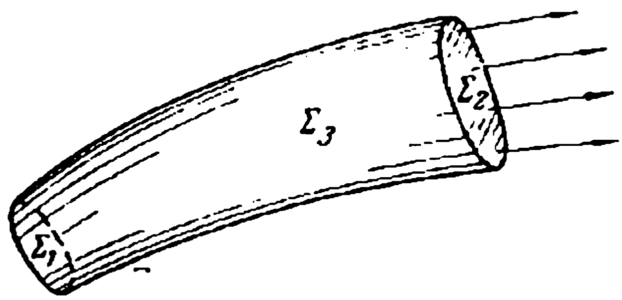


Fig. 6.7

These sections together with the lateral surface Σ_3 form a closed surface Σ. The field being solenoidal, i.e. $\operatorname{div} \mathbf{A} \equiv 0$, the Ostrogradsky theorem implies that

$$\iint_{\Sigma} A_n \, d\sigma = 0$$

But we have

$$\iint_{\Sigma} A_n \, d\sigma = \iint_{\Sigma_1} A_n \, d\sigma + \iint_{\Sigma_2} A_n \, d\sigma + \iint_{\Sigma_3} A_n \, d\sigma \qquad (6.16)$$

where each integral is taken over the outer side of the corresponding surface. The third summand on the right-hand side is equal to zero since, by the definition of a tube of a vector, the vector field **A** is directed perpendicularly to the normal to the surface Σ_3 at each point of this surface and therefore, on Σ_3, we have

$$A_n \equiv 0$$

If we now take the inner side of the section Σ_1, i.e. reverse the direction of its normal, and retain the outer side of the surface Σ_2 equality (6.16) turns into

$$\iint_{\Sigma_1} A_n \, d\sigma = \iint_{\Sigma_2} A_n \, d\sigma \qquad (6.17)$$

Hence, the flux of a solenoidal vector field **A** across every section of a vector tube has one and the same value. If the vector field **A** is interpreted as the velocity field of a flow of an incompressible liquid having neither sources nor sinks, relation (6.17) shows that the amount of liquid flowing through a cross section of a vector tube is the same for all the sections.

5. Equation of Continuity. As an application of the above notions, let us derive one of the basic equations of motion of fluid, the so-called *equation of continuity*. Let **A** be the field of velocities of a moving fluid. We shall suppose that the domain in which the flow is considered contains neither sources nor sinks, that is the fluid does not concentrate toward or expand from any point of the domain. But in contrast to our previous considerations, we shall not impose the condition that the fluid is incompressible and therefore the density ρ may depend not only on the coordinates x, y and z but also on time t.

Let us investigate the relationship between the velocity of the fluid and the variations of its density. For this purpose we pick out a closed volume Ω and find the increment ΔQ of the quantity of fluid (contained in Ω) corresponding to time period Δt by applying two different ways of computing ΔQ. Let $\rho(x, y, z, t)$ be the density of the fluid at moment t at an arbitrary point $(x, y, z) \in \Omega$. Then we obviously have

$$\Delta Q = \Delta t \iiint_{\Omega} \frac{\partial \rho}{\partial t}\, dv$$

On the other hand, the variation of the quantity of fluid contained within the volume Ω is equal to the flux of the fluid through the surface Σ (bounding the volume) multiplied by Δt. Therefore it is equal to $-\Delta t \iint_{\Sigma} (\rho \mathbf{A})_n\, d\sigma$ where n is the outer normal, the minus sign being taken because the quantity of fluid contained within a volume decreases when the velocity is directed outward. Transforming this surface integral by means of the Ostrogradsky theorem we obtain

$$\Delta Q = -\Delta t \iiint_{\Omega} \operatorname{div}(\rho \mathbf{A})\, dv$$

Now equating the two expressions for ΔQ and cancelling out Δt we receive

$$-\iiint_{\Omega} \operatorname{div}(\rho \mathbf{A})\, dv = \iiint_{\Omega} \frac{\partial \rho}{\partial t}\, dv$$

The last equality holding for any domain Ω, the integrands are equal, that is

$$\frac{\partial \rho}{\partial t} = -\operatorname{div}(\rho \mathbf{A}) \qquad (6.18)$$

We have thus derived an equation connecting the velocity and the density of a moving liquid for any region which contains neither sources nor sinks. Equation (6.18) is known as the equation of continuity.

Introducing the vector $\mathbf{J} = \rho\mathbf{A}$ (called the fluid-flux density vector) we can rewrite the equation of continuity in the form

$$\frac{\partial\rho}{\partial t} + \operatorname{div} \mathbf{J} = 0 \qquad (6.18')$$

6. Plane Fluid Flow. Ostrogradsky Theorem for Plane Field. Let us consider a plane vector field, that is one whose components have the form

$$P = P(x, y), \qquad Q = Q(x, y), \qquad R = 0 \qquad (6.19)$$

in an appropriately chosen Cartesian coordinate system (see § 2, Sec. 3). The field can be thought of as a velocity field of a fluid whose every particle moves in a plane parallel to the x, y-plane with a velocity independent of its distance to the latter (this is a so-called **plane fluid flow**). The divergence of such a field is equal to

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}$$

Let Ω be a right cylinder of unit altitude with base G (lying in the x, y-plane) and a lateral surface Σ (see Fig. 6.8). Now we put down the Ostrogradsky formula for the domain Ω. For this purpose

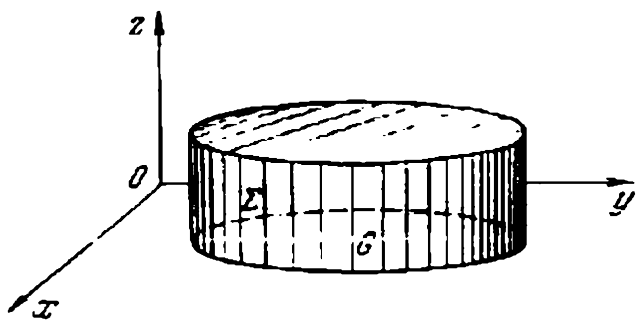


Fig. 6.8

we take into account that the numerical value of the triple integral of $\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}$ over Ω is equal to the double integral of this expression over the plane region G and that the flux of vector (6.19) across the surface Σ is equal to the line integral

$$\int\limits_L [P\cos(\mathbf{n}, x) + Q\cos(\mathbf{n}, y)]\, dl$$

where $\mathbf{n}$ is the normal (in the x, y-plane) to the contour L bounding the plane figure G. Besides, the fluxes through the upper and the lower bases of the cylinder Ω are equal to zero, the latter assertion being a consequence of the fact that vector (6.19) is perpendicular to the z-axis. It follows that the Ostrogradsky theorem for the plane field $\mathbf{A}$ and the cylindrical domain Ω is expressed by the relation

$$\int\limits_L [P\cos(\mathbf{n}, x) + Q\cos(\mathbf{n}, y)]\, dl = \iint\limits_G \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) dx\, dy \qquad (6.20)$$

Finally, we disregard the third coordinate z and consider (6.19) to be a vector field defined in the x, y-plane. Let us call the line integral

$$\int_L [P\cos(\mathbf{n}, x) + Q\cos(\mathbf{n}, y)]\,dl \tag{6.21}$$

the flux of vector field (6.19) through the contour L. Then formula (6.20) expressing the Ostrogradsky theorem for an arbitrary plane field $\mathbf{A}$ can be interpreted as follows: the double integral of the divergence of the plane field $\mathbf{A}$ over a domain G is equal to the flux of the vector $\mathbf{A}$ through the boundary L of the domain.

It can be easily verified that formula (6.20) is equivalent to Green's formula (4.45). Actually, if we denote by α the angle between the tangent to the curve L and the positive direction of the x-axis we can write

$$\cos(\mathbf{n}, x) = -\sin\alpha \quad \text{and} \quad \cos(\mathbf{n}, y) = \cos\alpha$$

Therefore integral (6.21) turns into

$$\int_L (Q\cos\alpha - P\sin\alpha)\,dl$$

or

$$\int_L Q\,dx - P\,dy$$

Now transforming the last line integral into the corresponding double integral by means of Green's theorem we arrive at formula (6.20).

The above argument can be reversed, i.e. if equality (6.20) has been established we can derive Green's formula from it.

Thus, Stokes' theorem and Ostrogradsky's theorem turn into Green's formula in the case of a plane field.

§ 4. CIRCULATION. ROTATION

1. Circulation of Vector Field. Let $\mathbf{A} = (P, Q, R)$ be a vector field and L be a smooth or piecewise smooth curve.

The line integral

$$\int_L P\,dx + Q\,dy + R\,dz$$

which can also be written as

$$\int_L A_\tau\,dl$$

where A_τ is the tangential component of the field $\mathbf{A}$ along the contour L (i.e. the projection of $\mathbf{A}$ on the tangent line to L) is referred to as the **circulation of the vector field A over the closed curve** L.

As was shown in § 2 of Chapter 4, if $\mathbf{A} = (P, Q, R)$ is a field of force, its circulation over a contour L is equal to the work of the force along the path L. For the fields of other nature the circulation may have another physical meaning.

2. Rotation of Vector Field. Stokes' Formula in Vector Notation. If L is a closed path the line integral

$$\int_L P\,dx + Q\,dy + R\,dz$$

can be transformed into the corresponding surface integral by applying Stokes' formula (5.41):

$$\oint_L P\,dx + Q\,dy + R\,dz = \iint_\Sigma \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy +$$
$$+ \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy\,dz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dz\,dx \tag{6.22}$$

The integral on the right-hand side of equality (6.22) is taken over an arbitrary surface Σ "pulled over" the contour L (i.e. a surface whose boundary is L). This integral can be interpreted as the flux, across the surface Σ, of the vector

$$\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k} \tag{6.23}$$

which is called the **rotation (curl) of the vector field A** and is denoted by the symbol rot $\mathbf{A}$ (or curl $\mathbf{A}$).

Thus, by definition,

$$\operatorname{rot}\mathbf{A} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k} \tag{6.24}$$

Using the notion of rotation we can rewrite Stokes' formula in the following form:

$$\oint_L A_\tau\,dl = \iint_\Sigma (\operatorname{rot}\mathbf{A})_n\,d\sigma \tag{6.25}$$

Thus, the circulation of a vector field $\mathbf{A}$ over a closed contour L is equal to the flux of the rotation of the vector field across an arbitrary surface (lying in the domain of definition of the field) whose boundary is L.

The above definition of the rotation of a vector field $\mathbf{A}$ involves not only the field itself but also a certain coordinate system x, y, z. But the vector rot $\mathbf{A}$ does not in fact depend on the choice of the

coordinate system and is uniquely determined by the field $\mathbf{A}$. To prove this let us write Stokes' formula (6.25) for a plane surface Σ and the contour L bounding this surface and then apply the mean value theorem to the surface integral on the right-hand side of equality (6.25). This results in

$$(\operatorname{rot} \mathbf{A}(M^*))_n = \frac{\oint\limits_L A_\tau\, dl}{\sigma}$$

where M^* is a point belonging to the surface Σ and σ is its area.* Let us now contract the surface Σ to a fixed point M so that the normal $\mathbf{n}$ to the surface all the time retains its direction invariable. In the limit we obtain

$$(\operatorname{rot} \mathbf{A}(M))_n = \lim_{\Sigma \to M} \frac{\oint\limits_L A_\tau\, dl}{\sigma} \tag{6.26}$$

Since the circulation of the vector $\mathbf{A}$ over the contour is independent of the choice of the coordinate system equality (6.26) implies that the projection of $\operatorname{rot} \mathbf{A}$ on the normal $\mathbf{n}$ does not depend on the choice. But the direction of the normal $\mathbf{n}$ can be chosen quite arbitrarily and thus the projection of the vector $\operatorname{rot} \mathbf{A}$ on any fixed axis is independent of the choice of the coordinate system. Consequently, the same is true for the vector $\operatorname{rot} \mathbf{A}$ itself.**

3. Symbolic Formula for Rotation. The expression of the rotation of a vector field $\mathbf{A} = (P, Q, R)$ can be conveniently written as the symbolic determinant

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} \tag{6.27}$$

where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are the unit vectors in the directions of the coordinate axes and the symbols $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$ are understood in the sense that the multiplication of such a symbol by a function means the differentiation with respect to the corresponding variable, e.g. $\frac{\partial}{\partial x} Q$ means $\frac{\partial Q}{\partial x}$.

---

* Here, as usual, we suppose that the first order partial derivatives of the components P, Q, R of the vector field $\mathbf{A}$ with respect to x, y, z are continuous.

** It is supposed that we take only right-handed coordinate systems. If we pass to a left-handed system (in which the positive direction of describing the boundary of a surface is such that the surface is always kept on the right) the direction of the vector $\operatorname{rot} \mathbf{A}$ changes to the opposite.

Indeed, (formally) expanding determinant (6.27) in minors of the first row we see that

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}$$

4. Physical Meaning of Rotation. The physical meaning of rotation can be elucidated in the following way. Let us again interpret a vector field $\mathbf{A}$ as the velocity field of a fluid flow. Imagine that we put, in the flow, an "infinitesimal turbine" whose plane vanes (parallel to the axis of the turbine) are placed round its circular periphery L (see Fig. 6.9). The fluid flow will make the turbine rotate about the

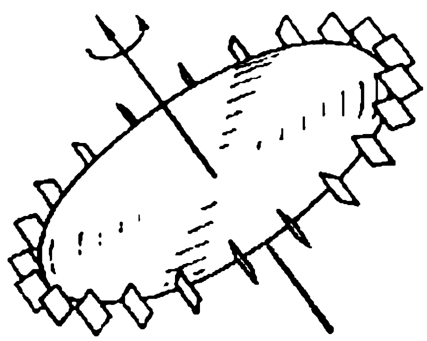


Fig. 6.9

axis with an angular velocity which is, generally speaking, dependent on the direction of the axis.

It appears natural to assume that the magnitude v of the linear velocity of each point of the circle L is equal to the mean value of the projections of the vector $\mathbf{A}$ on the corresponding tangents to L that is

$$v = \frac{1}{2\pi R} \oint_L A_\tau \, dl \tag{6.28}$$

By Stokes' formula (6.25), line integral (6.28) can be transformed into the surface integral

$$\frac{1}{2\pi R} \iint_\Sigma (\operatorname{rot} \mathbf{A})_n \, d\sigma \tag{6.29}$$

taken over the surface Σ of the turbine. The turbine being considered infinitesimal, we can write the integral $\iint_\Sigma (\operatorname{rot} \mathbf{A})_n \, d\sigma$ as the product of the area of the turbine by the value of $(\operatorname{rot} \mathbf{A})_n$ at its centre i.e. in the form

$$\pi R^2 (\operatorname{rot} \mathbf{A})_n$$

Relation (6.28) then takes the form

$$v = \frac{R}{2}(\mathrm{rot}\,\mathbf{A})_n$$

As is known, the projection of a vector assumes its maximal value (equal to the modulus of the vector) when the axis the vector is projected on is parallel to the vector and has the same direction. Therefore, if we place the axis of the turbine so that its speed v is maximal (i.e. so that the direction of the axis coincides with the direction of rot A) we obtain

$$v_{max} = \frac{R}{2}|\,\mathrm{rot}\,\mathbf{A}\,|$$

or

$$|\,\mathrm{rot}\,\mathbf{A}\,| = \frac{2v_{max}}{R}$$

But $\frac{v}{R}$ is equal to the angular speed ω of the turbine and hence we have arrived at the following result: if the turbine is placed in the flow so that its speed of rotation becomes maximal its angular speed is equal to $\frac{1}{2}$ | rot A | and the direction of its axis coincides with the direction of the vector rot A.

Consequently, the vector rot A characterizes "the rotational component" of the velocity field and is equal to the doubled angular velocity of rotation of an infinitesimal particle of the fluid.

Examples

1. Consider the vector field A with the components

$$P = -y\omega, \quad Q = x\omega, \quad R = 0$$

The field can be regarded as the velocity field corresponding to the rotation of the entire space (filled with a fluid) about the z-axis with angular velocity $\boldsymbol{\omega} = (0, 0, \omega)$. It can be easily verified that the rotation of the vector field is equal to $2\omega\mathbf{k}$, that is rotA ($\mathbf{A} = (P, Q, R)$) is directed along the axis of revolution and its magnitude is twice the angular speed ω (see Fig. 6.10).

The physical significance of this result is that every particle (of a fluid which is in a rotary motion about the z-axis) passing through a point (x, y, z) at moment t simultaneously takes part in two motions, namely in the instantaneous transient motion with velocity $\mathbf{v} = (-y\omega, x\omega, 0)$ and in the instantaneous rotary motion. It is obvious that the instantaneous angular velocity of rotation of every particle coincides with the angular velocity $\boldsymbol{\omega}$ of the macroscopic motion of the fluid as a whole. Hence, the field of the rotation rotA is also constant and equal to $2\boldsymbol{\omega}$ and the fluid can be thought of as being entirely composed of infinitesimal curls (vortices).

2. Consider a liquid flowing in a constant direction with a constant velocity, i.e. suppose that the functions P, Q and R are identically constant. In this case we have rot $\mathbf{A} \equiv 0$.

3. Let $P = y$, $Q = 0$ and $R = 0$. Then

$$\text{rot } \mathbf{A} = -\mathbf{k}$$

Here the rotation is different from zero at each point although all the vector lines are straight lines parallel to the y, z-plane (see Fig. 6.11). One may think that this result contradicts the assertion that rot $\mathbf{A}$ characterizes the "rotational component" of the field $\mathbf{A}$. But in fact this is not so: in this example the "rotational component"

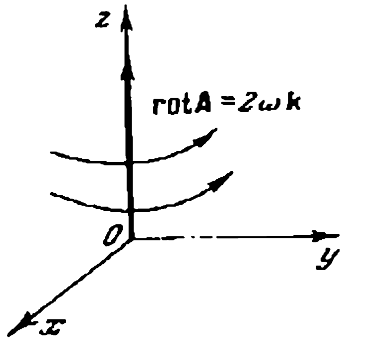


Fig. 6.10

is due not to the twist of the vector lines (which are rectilinear here) but to the variation of the velocity which is dependent on the distance from the x, z-plane. If we place an infinitesimal turbine in this

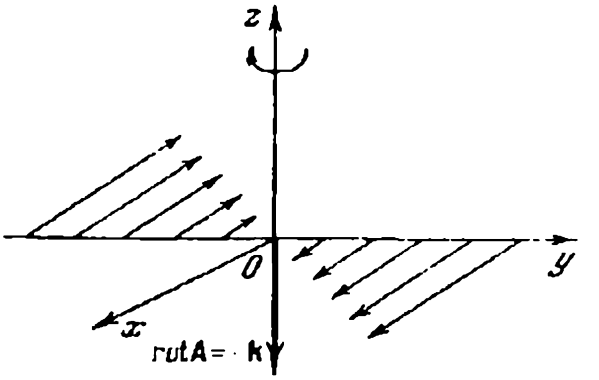


Fig. 6.11

flow whose velocity is equal to $(y, 0, 0)$ at each point (x, y, z) it will not apparently be in a state of rest unless its axis of rotation is perpendicular to the z-axis.

4. Let us take the vector field $\mathbf{A}$ with the components

$$P = \frac{x}{\sqrt{x^2+y^2}}, \qquad Q = \frac{-y}{\sqrt{x^2+y^2}}, \qquad R = 0 \tag{6.30}$$

This field can be regarded as the velocity field of a fluid whose particles move, in the x, y-plane, along the hyperbolas $xy = C$ (Fig. 6.12) in such a way that the magnitude of the velocity is equal to unity at each point. Computing the divergence and the rotation of the

field we find

$$\operatorname{div} \mathbf{A} = \frac{\partial}{\partial x}\left(\frac{x}{\sqrt{x^2+y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{-y}{\sqrt{x^2+y^2}}\right) = \frac{y^2-x^2}{(x^2+y^2)^{3/2}}$$

and

$$\operatorname{rot} \mathbf{A} = \left[\frac{\partial}{\partial x}\left(\frac{-y}{\sqrt{x^2+y^2}}\right) - \frac{\partial}{\partial y}\left(\frac{x}{\sqrt{x^2+y^2}}\right)\right]\mathbf{k} = -\frac{2xy}{(x^2+y^2)^{3/2}}\mathbf{k}$$

The divergence is positive here for $|y| > |x|$ and negative for $|y| < |x|$. From the physical point of view, this means that the motion of an incompressible fluid described by field (6.30) is only possible if there are sources in the regions where $|y| > |x|$ and sinks where $|y| < |x|$. The rotation of field (6.30) is directed along the z-axis at each point, this being so for every plane field parallel to the x, y-plane. In the case of field (6.30) the vector rot **A** goes along the positive direction of the z-axis in the second and fourth quadrants and in the negative direction in the first and second quadrants. Both divergence and rotation of field (6.30) tend to zero as $x^2 + y^2 \to \infty$, that is when the distance from the origin of coordinates increases.

5. **More on Potential and Solenoidal Fields.** The notion of rotation discussed in Sec. 2 is directly related to the definitions of a potential and of a solenoidal field introduced above.

We have defined a potential vector field as one representable in the form of the gradient of a scalar field. As was shown, a vector field $\mathbf{A} = (P, Q, R)$ is potential if and only if its components satisfy* the conditions

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y} \quad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z}$$

But these three conditions mean that all the three components of the rotation of the field **A** are identically equal to zero.

Thus, we can state that for a vector field **A** to be potential it is necessary and sufficient that the condition**

$$\operatorname{rot} \mathbf{A} \equiv 0$$

be fulfilled. Hence, we always have rot grad $U = 0$ for any function U.

The notion of a solenoidal field introduced in § 2 is also connected with the notion of rotation. Indeed, the direct differentiation shows

* As before, we suppose that the functions P, Q and R are continuously differentiable and that the domain in which the field is considered is simply connected.

** In this case we suppose that the components of the vector field **A** defined in a simply connected domain possess the continuous partial derivatives of the first and second orders with respect to x, y and z.

that, for every vector field **A**, we have the relation

$$\operatorname{div}(\operatorname{rot}\mathbf{A})=\frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)+\frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)=0$$

and hence any vector field representable as the rotation of another vector field is always solenoidal.

It can be proved that, conversely, every solenoidal field can be represented in the form of the rotation of a vector field. In other words, for every vector field **A** satisfying the condition $\operatorname{div}\mathbf{A}=0$

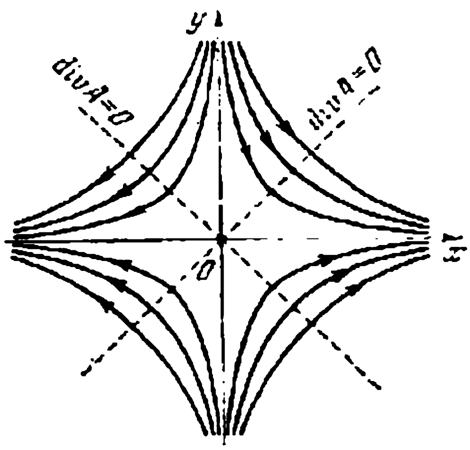


Fig. 6.12

there exists a field **B** such that $\mathbf{A}=\operatorname{rot}\mathbf{B}$. Such a vector field **B** is not uniquely determined by the field **A** but only to within a summand of the form $\operatorname{grad}U$, U being an arbitrary function.

If $\mathbf{A}=\operatorname{rot}\mathbf{B}$ the field **B** is termed a **vector potential of the field A.**

Although the potential and the solenoidal fields do not exhaust all possible vector fields it can be shown that an arbitrary vector field can be represented as a combination of the fields of these two types. More precisely, it can be proved that every vector field **A** can be represented in the form

$$\mathbf{A}=\mathbf{B}+\mathbf{C}$$

where the field **B** is potential and **C** is solenoidal.

§ 5. HAMILTONIAN OPERATOR

1. Symbolic Vector ∇. In § 1 we introduced the notion of the gradient of a scalar field. The process of obtaining the vector field $\operatorname{grad}U$ from a given scalar field U can be regarded as an operation which is in many respects similar to differentiation with the only difference that the latter transforms a scalar into a new scalar whereas the former operation applied to a scalar yields a vector.

The operation of passing from U to $\operatorname{grad}U$ is usually designated by the symbol ∇ introduced by Hamilton*. This symbol is read

* Hamilton, William Rowan (1805-1865), an Irish mathematician.

"nabla"* (or "del") and is called the **Hamiltonian operator.** Thus, we have, by definition,

$$\nabla U = \text{grad } U$$

It is convenient to interpret the operator ∇ as a symbolic vector with the components $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$**:

$$\nabla = \frac{\partial}{\partial x}\,\mathbf{i} + \frac{\partial}{\partial y}\,\mathbf{j} + \frac{\partial}{\partial z}\,\mathbf{k}$$

The application of this operator to a scalar function U can be performed as formal multiplication of the "vector" ∇ by the scalar U:

$$\nabla U = \text{grad } U = \left(\frac{\partial}{\partial x}\,\mathbf{i} + \frac{\partial}{\partial y}\,\mathbf{j} + \frac{\partial}{\partial z}\,\mathbf{k}\right)U = \mathbf{i}\,\frac{\partial U}{\partial x} + \mathbf{j}\,\frac{\partial U}{\partial y} + \mathbf{k}\,\frac{\partial U}{\partial z}$$

The vector ∇ can also be conveniently used for writing formulas involving many other operations of vector analysis. For instance, if $\mathbf{A} = (P, Q, R)$ then

$$\text{div } \mathbf{A} = \frac{\partial}{\partial x}P + \frac{\partial}{\partial y}Q + \frac{\partial}{\partial z}R = (\nabla, \mathbf{A})$$

that is the divergence of a vector field $\mathbf{A}$ is the (formal) scalar product of the symbolic vector ∇ by the vector $\mathbf{A}$. Similarly, we have

$$\text{rot } \mathbf{A} = \left(\frac{\partial}{\partial y}R - \frac{\partial}{\partial z}Q\right)\mathbf{i} + \left(\frac{\partial}{\partial z}P - \frac{\partial}{\partial x}R\right)\mathbf{j} + \\ + \left(\frac{\partial}{\partial x}Q - \frac{\partial}{\partial y}P\right)\mathbf{k} = [\nabla, \mathbf{A}]$$

i.e. the rotation of a vector field $\mathbf{A}$ is the (formal) vector product of the vector ∇ by the vector $\mathbf{A}$.

2. Operations with Vector ∇. The expedience of the introduction of the symbolic vector ∇ lies in the fact that it enables us to deduce and write various formulas of vector analysis in an abbreviated and visual form. There are some simple examples below.

In Sec. 5 we showed, by means of the direct calculations, that

$$\text{rot grad } U = 0$$

and

$$\text{div rot } \mathbf{A} = 0$$

---

* This term, also introduced by Hamilton, originates from the name nabla (Greek ναβλα) of an ancient musical instrument of triangular form.

** When writing the rotation as a symbolic determinant, we have already seen that it is convenient to think of the operation of differentiation as the (formal) multiplication of the symbol $\frac{\partial}{\partial x}$ $\left(\text{or } \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \text{ etc.}\right)$ by the function whose derivative is computed.

These identities can be rewritten with the help of the vector ∇ as two relations of the form

$$[\nabla, \nabla U] = 0$$

and

$$(\nabla, \nabla, \mathbf{A}) = 0$$

The left-hand side of the former is the "vector product" of two "vectors" ∇ and ∇U which only differ from each other in the scalar factor U, and we have the "triple scalar product" involving two equal vectors on the left-hand side of the latter relation. Therefore these expressions turn out to be equal to zero in accordance with the general rules of vector algebra.

We can directly verify that most of the basic operations performed on the ordinary vectors can be extended to the symbolic vector ∇ which makes it possible to derive formulas of vector analysis by applying the rules of vector algebra to expressions involving ∇.

But it should be noted that there is no complete analogy between the symbolic vector ∇ and "true" vectors. Namely, only the formulas (containing the vector ∇) which do not involve the application of the operator ∇ to products of variable quantities (scalar or vector) are completely analogous to the corresponding formulas of vector algebra. But if an expression contains a product of two or more variable quantities to which the operator ∇ should be applied we cannot follow the ordinary rules of vector algebra. To establish the corresponding rules for application of the symbolic operator ∇ to such expressions we shall consider some typical examples.

1. Let $U = U(x, y, z)$ be a scalar field and $\mathbf{A} = \mathbf{A}(x, y, z)$ a vector field. We shall find $\operatorname{div}(U\mathbf{A})$, i.e. $(\nabla, U\mathbf{A})$. The application of the vector operator ∇ reduces to performing the operations of differentiation involved in it. But, as is well known, when differentiating a product we can differentiate the first factor considering the others to be constant, then differentiate the second factor as if the other factors were constant and so on and then take the sum of the expressions thus obtained.

Let us mark by the sign "$\downarrow$" the factor to which the operator ∇ should be applied. Then, as it can be easily verified, the expression for $\operatorname{div}(U\mathbf{A})$ can be written as

$$(\nabla, U\mathbf{A}) = (\nabla, \overset{\downarrow}{U}\mathbf{A}) + (\nabla, U\overset{\downarrow}{\mathbf{A}})$$

The factors, in each summand, that are not subjected to the operator ∇ can be taken outside the sign ∇. Consequently we obtain

$$(\nabla, U\mathbf{A}) = (\nabla, \overset{\downarrow}{U}\mathbf{A}) + (\nabla, U\overset{\downarrow}{\mathbf{A}}) = (\nabla U, \mathbf{A}) + U(\nabla, \mathbf{A})$$

which can be written down in the ordinary notation as

$$\operatorname{div}(U\mathbf{A}) = (\mathbf{A}, \operatorname{grad} U) + U \operatorname{div} \mathbf{A}$$

2. Consider the expression

$$\operatorname{grad}(UV)$$

which can be written symbolically in the form

$$\nabla UV$$

Following the above rule we can write

$$\nabla UV = \nabla \check{U}V + \nabla U\check{V} = V\nabla U + U\nabla V$$

or, in the ordinary notation,

$$\operatorname{grad}(UV) = V \operatorname{grad} U + U \operatorname{grad} V$$

These examples enable us to formulate the rules according to which the operator ∇ should be applied: in the expressions containing a single variable quantity the symbol ∇ can be operated on as an ordinary vector and in the expressions involving products of several variable quantities the operator should be applied in accord with the rule for differentiating a product. Finally, the application of ∇ to a sum of any summands is performed termwise, i.e. ∇ is separately applied to each summand and the results are then added up.

In conclusion, we present a list of formulas connecting the operations of computing the gradient, the rotation and the divergence with basic operations of vector algebra:

1. $\operatorname{div}(U\mathbf{A}) = (\mathbf{A}, \operatorname{grad} U) + U \operatorname{div} \mathbf{A}$,
2. $\operatorname{grad}(UV) = V \operatorname{grad} U + U \operatorname{grad} V$,
3. $\operatorname{rot}(U\mathbf{A}) = U \operatorname{rot} \mathbf{A} + [\operatorname{grad} U, \mathbf{A}]$,
4. $\operatorname{div}[\mathbf{A}, \mathbf{B}] = (\mathbf{B}, \operatorname{rot} \mathbf{A}) - (\mathbf{A}, \operatorname{rot} \mathbf{B})$,
5. $\operatorname{rot}[\mathbf{A}, \mathbf{B}] = (\mathbf{B}, \nabla)\mathbf{A} - (\mathbf{A}, \nabla)\mathbf{B} + \mathbf{A} \operatorname{div} \mathbf{B} - \mathbf{B} \operatorname{div} \mathbf{A}$,
6. $\operatorname{grad}(\mathbf{A}, \mathbf{B}) = (\mathbf{B}, \nabla)\mathbf{A} + (\mathbf{A}, \nabla)\mathbf{B} + [\mathbf{B}, \operatorname{rot} \mathbf{A}] +$
$+ [\mathbf{A}, \operatorname{rot} \mathbf{B}]$.

In particular, putting $\mathbf{A} = \mathbf{B}$ in the last formula we derive

$$\operatorname{grad} \frac{A^2}{2} = (\mathbf{A}, \nabla)\mathbf{A} + [\mathbf{A}, \operatorname{rot} \mathbf{A}]$$

The first two of these formulas have already been deduced. The other can be obtained by applying the operator ∇ in accordance with the above rules and ordinary formulas of vector algebra. In particular, to find the expression for $\operatorname{rot}[\mathbf{A}, \mathbf{B}]$ we symbolically write it in the form

$$[\nabla, [\mathbf{A}, \mathbf{B}]]$$

and apply the well known formula* for transforming a triple vector product:

$$[\mathbf{a}, [\mathbf{b}, \mathbf{c}]] = \mathbf{b}(\mathbf{a}, \mathbf{c}) - \mathbf{c}(\mathbf{a}, \mathbf{b})$$

---

* For a triple vector product of the form $[[\mathbf{a}, \mathbf{b}], \mathbf{c}]$ the corresponding formula is written as

$$[[\mathbf{a}, \mathbf{b}], \mathbf{c}] = \mathbf{b}(\mathbf{a}, \mathbf{c}) - \mathbf{a}(\mathbf{b}, \mathbf{c})$$

An expression of the form $(\mathbf{A}, \nabla)\,\mathbf{B}$ encountered in formulas 5 and 6 is understood as the vector quantity

$$\left(A_x \frac{\partial B_x}{\partial x} + A_y \frac{\partial B_x}{\partial y} + A_z \frac{\partial B_x}{\partial z},\ A_x \frac{\partial B_y}{\partial x} + A_y \frac{\partial B_y}{\partial y} + A_z \frac{\partial B_y}{\partial z},\right.$$
$$\left. A_x \frac{\partial B_z}{\partial x} + A_y \frac{\partial B_z}{\partial y} + A_z \frac{\partial B_z}{\partial z}\right)$$

(where A_x, A_y, A_z and B_x, B_y, B_z are the components of the vectors $\mathbf{A} = (A_x, A_y, A_z)$ and $\mathbf{B} = (B_x, B_y, B_z)$) which can be regarded as the result of applying the "scalar" operation

$$(\mathbf{A}, \nabla) = A_x \frac{\partial}{\partial x} + A_y \frac{\partial}{\partial y} + A_z \frac{\partial}{\partial z}$$

to each component of the vector $\mathbf{B}$.*

§ 6. REPEATED OPERATIONS INVOLVING ∇. LAPLACIAN OPERATOR

1. Repeated Differentiation. In §§ 3-5 we introduced the notions of gradient, divergence and rotation. In applications of vector analysis we deal not only with these basic operations but also with their combinations. The most often used operations of this kind are those containing second derivatives of the fields.

We can compose nine different combinations of the symbols grad, rot and div involving second derivatives but some of them are senseless. For instance, such is the operation

$$\text{rot div}$$

which cannot be applied either to a scalar field or to a vector field because the notion of the divergence has been introduced for the vector fields $\mathbf{A}$ and the expression $\operatorname{div} \mathbf{A}$ is always a scalar quantity whereas the rotation was only defined for the vector fields and thus it is senseless to speak about the rotation of a divergence. All possible combinations are given in the table below, the senseless combinations being marked by shading the corresponding positions of the table (see the next page).

We see that, among these operations, there are only two that can be applied to a scalar field U, namely

$$\text{rot grad}\ U$$

and

$$\text{div grad}\ U$$

---

* In the formula for $(\mathbf{A}, \nabla\ \mathbf{B})$ we have denoted the components of the vectors $\mathbf{A}$ and $\mathbf{B}$ by the same letters A and B with subscripts x, y and z (designating the coordinate axes) which makes the formula look symmetric with respect to the letters. In what follows we shall also use this notation.

Operation	Scalar field U	Vector field $\mathbf{A}$	
	grad	div	rot
grad		grad div $\mathbf{A}$	
div	div grad U		div rot $\mathbf{A} \equiv 0$
rot	rot grad $U \equiv 0$		rot rot $\mathbf{A}$

The former is the rotation of the potential vector field grad U and, as was shown, rot grad U is always identically equal to zero. The latter expression div grad U is not equal to zero in the general case. It is called the **Laplacian operator*** and denoted by ΔU. Taking advantage of the expressions for the gradient and rotation in Cartesian coordinates (see formulas (6.5) and (6.23)) we derive

$$\Delta U = \operatorname{div} \operatorname{grad} U = \operatorname{div}\left(\frac{\partial U}{\partial x}\mathbf{i} + \frac{\partial U}{\partial y}\mathbf{j} + \frac{\partial U}{\partial z}\mathbf{k}\right) = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2}$$

The divergence and the gradient being independent of the particular choice of the coordinate system, the quantity ΔU is thus completely specified by the field U and does not depend on the coordinate system. Later we shall dwell in more detail on Laplace's operator.

The Laplacian operator Δ can be considered to be the (formal) scalar square of the symbolic vector ∇. In fact, we have

$$(\nabla, \nabla) = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 + \left(\frac{\partial}{\partial z}\right)^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

and hence

$$(\nabla, \nabla) U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = \Delta U, \quad \text{i.e.} \ (\nabla, \nabla) = \Delta$$

It is sometimes necessary to apply the operator Δ not to a scalar quantity but to a vector. If

$$\mathbf{A} = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$$

* Named after P. S. Laplace (1749-1827), a prominent French mathematician and astronomer. More precisely, the term "Laplacian operator" is used not for the expression ΔU (which is the result of applying the operator to a scalar field U) but for the symbol $\Delta \equiv$ div grad.

then, by definition, $\Delta\mathbf{A}$ is understood as the vector

$$\Delta A_x\mathbf{i} + \Delta A_y\mathbf{j} + \Delta A_z\mathbf{k}$$

As will be shown later, this quantity does not in fact depend on the choice of the coordinate system and is completely determined by the vector field $\mathbf{A}$.*

Let us now proceed to study the operations involving repeated differentiation which make sense for the vector fields. There are three such operations in the above table, namely

$$\operatorname{grad} \operatorname{div} \mathbf{A}$$

$$\operatorname{rot} \operatorname{rot} \mathbf{A}$$

$$\operatorname{div} \operatorname{rot} \mathbf{A}$$

The expression div rot $\mathbf{A}$ was encountered in § 4 when we deduced the conditions for a field to be solenoidal. As was shown, we always have

$$\operatorname{div} \operatorname{rot} \mathbf{A} \equiv 0$$

On the contrary, the expressions grad div $\mathbf{A}$ and rot rot $\mathbf{A}$ may be nonzero. They are widely applied in various problems of mechanics and electrodynamics.

Let us derive a formula connecting these quantities. For this purpose we first consider the expression

$$\operatorname{rot} \operatorname{rot} \mathbf{A}$$

which can be written, in symbolic form, as

$$[\nabla, [\nabla, \mathbf{A}]]$$

Applying the formula for a triple vector product given in § 5, Sec. 2, we obtain

$$[\nabla, [\nabla, \mathbf{A}]] = \nabla (\nabla, \mathbf{A}) - (\nabla, \nabla) \mathbf{A}$$

that is

$$\operatorname{rot} \operatorname{rot} \mathbf{A} = \operatorname{grad} \operatorname{div} \mathbf{A} - \Delta\mathbf{A} \qquad (6.31)$$

In particular, it follows that the quantity $\Delta\mathbf{A}$ defined above as the result of application of the Laplacian operator to each component of the vector function $\mathbf{A}$ is in fact independent of the choice of the coordinate system because this is the case for the quantities rot rot $\mathbf{A}$ and grad div $\mathbf{A}$.

Expression (6.31) involving only one variable quantity, we have applied here the ordinary rules of vector algebra in operating on the symbol ∇. We suggest that the reader directly derive equality (6.31) without resorting to the symbol ∇ and compare the calculations

---

* It should be noted that the original definition of Δ as Δ = div grad only applies when we consider an expression of the form ΔU where U is a scalar.

with the above because this will show the advantage of introducing the Hamiltonian operator.

2. Heat Conductivity Equation. To illustrate the possibility of application of the concepts of vector analysis let us derive the equation for a temperature field inside a physical body subjected to heating. Denote by $U(x, y, z, t)$ the temperature at an arbitrary point (x, y, z) of the body at moment t. Let us (mentally) pick out, within the body, a volume Ω bounded by a surface Σ and compute, by means of two different ways, the variation of the amount of heat, contained in the volume, during an infinitesimal time period dt. In every infinitesimal element of volume dv taken inside the body the increment of the temperature corresponding to the time interval dt is equal to $\frac{\partial U}{\partial t}dt$, the mass of the element being equal to $\rho\, dv$ where ρ is the mass density. Consequently, the variation of the quantity of heat in the volume element dv is expressed as

$$c\frac{\partial U}{\partial t}dt\rho\, dv$$

where c is the *specific heat*, the quantities c and ρ being considered constant. Accordingly, the variation of the quantity of heat, contained in the whole volume Ω, during time dt is equal to

$$dQ = dt\iiint\limits_{\Omega}\frac{\partial U}{\partial t}c\rho\, dv$$

On the other hand, the quantity dQ can be found by computing the amount of heat passing through the surface Σ, bounding the volume Ω, during the time period from t to $t+dt$. By Sec. 1 of § 2, the amount of heat passing during the time interval dt through an elementary area $d\sigma$ (in the direction of the outer normal to Σ) is equal to

$$-\,dt\,k\,(\operatorname{grad} U)_n\, d\sigma$$

and consequently the resultant increment of quantity of heat contained within the volume Ω bounded by the surface Σ is given by the surface integral

$$dQ = dt\iint\limits_{\Sigma}k\,(\operatorname{grad} U)_n\, d\sigma$$

Transforming this integral into the corresponding volume integral, by applying the Ostrogradsky theorem, we receive

$$dt\iint\limits_{\Sigma}k\,(\operatorname{grad} U)_n\, d\sigma = dt\iiint\limits_{\Omega}k\operatorname{div}(\operatorname{grad} U)\, dv = dt\iiint\limits_{\Omega}k\Delta U\, dv$$

Now equating the two expressions obtained for dQ and cancelling out dt we deduce the relation

$$\iiint_{\Omega} \frac{\partial U}{\partial t} c\rho \, dv = \iiint_{\Omega} k \Delta U \, dv$$

which is valid for any spatial region Ω taken within the body. It follows that the integrands on the left-hand and right-hand sides are equal, that is

$$\frac{\partial U}{\partial t} = a^2 \Delta U \qquad \left(a^2 = \frac{k}{c\rho}\right) \tag{6.32}$$

We have thus derived the equation, called the **heat conductivity equation**, for the function U describing distribution of temperature in a heat conductor.

3. Stationary Distribution of Temperature. Harmonic Functions We have shown that a distribution of temperature inside a physical body must satisfy equation (6.32). In a particular case it may turn out that the body is in a *state of thermal equilibrium*, i.e. its temperature does not vary either on the boundary or in the interior of the body. Then we must have $\frac{\partial U}{\partial t} \equiv 0$, and equation (6.32) takes the form

$$\Delta U = 0$$

or, in Cartesian coordinates,

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0$$

A state of thermal equilibrium can be thought of as follows. Suppose that a constant temperature (independent of time) is kept at each point of the boundary of the body (but the temperature may vary in the general case, from point to point on the boundary). Then after a sufficiently large time period has elapsed (strictly speaking an infinite time period) there will be a stationary, time-independent, temperature distribution. Such a distribution retains a constant value (which may be different for different points) at each point of the body. Apparently, such an equilibrium temperature distribution is uniquely determined by the temperature conditions on the boundary of the body.

The equation

$$\Delta U = 0$$

is known as **Laplace's equation**. There are various stationary processes distinct from heat conduction which are also described by this equation. For instance, these are an equilibrium distribution of electric charge over the surface of a conductor, a stationary fluid flow in a closed region and the like. A scalar function $U(x, y, z)$ satisfying the equation $\Delta U = 0$ is called a **harmonic function**

(field). Thus, a stationary distribution of temperature within a physical body is described by a harmonic function.

The function

$$\frac{k}{r} \quad \left(r=\sqrt{x^2+y^2+z^2}, \quad k=\text{const}\right)$$

is one of the most important examples of a harmonic function. It can be interpreted as the potential function of the field of gravitation (the electrostatic field) produced by a material point (point charge) placed at the origin of coordinates. Let us verify that this is in fact a harmonic function everywhere except the origin at which it is not defined. Indeed, we have

$$\frac{\partial}{\partial x}\,\frac{k}{r}=-\frac{kx}{r^3}, \quad \frac{\partial^2}{\partial x^2}\,\frac{k}{r}=-k\,\frac{r^3-3r^2x}{r^6}=k\,\frac{3x^2-r^2}{r^5}$$

and, similarly,

$$\frac{\partial^2}{\partial y^2}\,\frac{k}{r}=k\,\frac{3y^2-r^2}{r^5}, \quad \frac{\partial^2}{\partial z^2}\,\frac{k}{r}=k\,\frac{3z^2-r^2}{r^5}$$

whence

$$\Delta\left(\frac{k}{r}\right)=0 \quad (r\neq 0)$$

A function of the form $\frac{k}{|\mathbf{r}-\mathbf{r}_0|}$ is also harmonic for each fixed $\mathbf{r}_0$ at every point whose radius vector $\mathbf{r}$ is distinct from $\mathbf{r}_0$. Consequently, any linear combination of the form

$$\sum_{i=1}^{n}\frac{k_i}{|\mathbf{r}-\mathbf{r}_i|}$$

is a harmonic function (for $\mathbf{r}\neq\mathbf{r}_i$, $i=1, 2, \ldots, n$) which can be interpreted as the potential function of the field of gravitation (the electrostatic field) produced by a system of material points (point charges) located at the points with radius vectors $\mathbf{r}_i$ ($i=1, 2, \ldots, n$). The field of gravitation produced by a mass continuously distributed in space with volume density $\mu(x, y, z)$ can be obtained by performing the passage to the limit in the above expression from the system of mass points to a continuous mass distribution. The rigorous mathematical justification of such a passage to the limit, which looks quite natural from the point of view of physics, would involve integrals dependent on parameters (see Chapter 10). We shall not present such a proof here because these questions fall outside the framework of our course and are considered in the textbooks on mathematical physics.

Exercises

1. Write down the expressions for the potential function of the field of gravitation produced by a mass continuously distributed in space with density $\mu(x, y, z)$.

2. Find the potential of the electrostatic field produced by an infinite thread carrying a uniformly distributed charge. Is this potential a harmonic function?

§ 7. EXPRESSING FIELD OPERATIONS IN CURVILINEAR ORTHOGONAL COORDINATES

1. Statement of the Problem. Such quantities as the gradient, divergence, rotation etc. of a field are widely applied in various problems of theoretical and mathematical physics. In many cases it appears necessary to express these quantities not only in Cartesian coordinates, as has been done in §§ 3-6, but also in curvilinear coordinate systems. For instance, if we deal with a spherically symmetric field, i.e. if, at every point in space, there is a scalar or vector quantity dependent only on the distance from the point to the origin of coordinates, it is clear that all the formulas related to such a field are essentially simplified if we write them in spherical coordinates instead of Cartesian ones. Of course, in various problems some other coordinate systems may prove to be more convenient.

Here we shall derive the formulas for the gradient, divergence, rotation and Laplace's operator in general curvilinear orthogonal coordinates and also consider the special cases of spherical and cylindrical coordinates.

2. Curvilinear Orthogonal Coordinates in Space. Suppose we are given a curvilinear coordinate system* q_1, q_2, q_3 in space and let the formulas connecting q_1, q_2 and q_3 with the Cartesian coordinates x, y and z be of the form

$$x = x(q_1, q_2, q_3), \quad y = y(q_1, q_2, q_3), \quad z = z(q_1, q_2, q_3) \qquad (6.33)$$

We shall confine ourselves to the simplest case of *orthogonal coordinate systems* which are particularly important for practical applications. A curvilinear coordinate system is said to be orthogonal if at each point the three coordinate curves passing through the point form right angles with each other, i.e. the tangents to

---

* The notion of curvilinear coordinates was discussed in Chapter 2. We shall suppose that the functions expressing curvilinear coordinates in terms of Cartesian ones satisfy the conditions enumerated in § 4 of Chapter 2. We shall change the notation introduced in Chapter 2 and designate the curvilinear coordinates by the single letter q with the subscripts 1, 2 and 3 so that the formulas take symmetric form. Similarly, for the same reason, the components of a vector field $\mathbf{A}$ (with respect to the coordinate system in question) will be labelled by A_1, A_2 and A_3.

the curves are mutually orthogonal at the point. In particular, the spherical and the cylindrical coordinate systems in space which are widely used in various problems also possess the orthogonality property.

First of all, let us find the expressions for the elements of length, area and volume in orthogonal coordinates. For this purpose we consider an infinitesimal curvilinear parallelepiped cut out of space by three pairs of coordinate surfaces corresponding, respectively, to the values q_1, q_1+dq_1, q_2, q_2+dq_2 and q_3, q_3+dq_3 of the coordinates q_1, q_2 and q_3 (see Fig. 6.13).

Consider the edge MM_1 of the parallelepiped. The point M has the curvilinear coordinates (q_1, q_2, q_3) and the point M_1 has the coordinates (q_1+dq_1, q_2, q_3). Denoting the Cartesian coordinates of the point M as x, y, z and those of the point M_1 as $x+dx$, $y+dy$, $z+dz$ we can write the expression

$$\sqrt{dx^2+dy^2+dz^2}$$

for the length dl_1 of the vector $\overrightarrow{MM_1}$. The coordinates x, y and z are functions of a single parameter q_1 along the edge MM_1 since q_2 and q_3 are constant on MM_1. Consequently, in this case we have

$$dx=\frac{\partial x}{\partial q_1}dq_1, \quad dy=\frac{\partial y}{\partial q_1}dq_1, \quad dz=\frac{\partial z}{\partial q_1}dq_1$$

and

$$dl_1=\sqrt{\left(\frac{\partial x}{\partial q_1}\right)^2+\left(\frac{\partial y}{\partial q_1}\right)^2+\left(\frac{\partial z}{\partial q_1}\right)^2}\,dq_1$$

Similarly, for the lengths dl_2 and dl_3 of the edges MM_2 and MM_3 we derive the expressions

$$dl_2=\sqrt{\left(\frac{\partial x}{\partial q_2}\right)^2+\left(\frac{\partial y}{\partial q_2}\right)^2+\left(\frac{\partial z}{\partial q_2}\right)^2}\,dq_2$$

and

$$dl_3=\sqrt{\left(\frac{\partial x}{\partial q_3}\right)^2+\left(\frac{\partial y}{\partial q_3}\right)^2+\left(\frac{\partial z}{\partial q_3}\right)^2}\,dq_3$$

Introducing the notation

$$\left.\begin{aligned} H_1&=\sqrt{\left(\frac{\partial x}{\partial q_1}\right)^2+\left(\frac{\partial y}{\partial q_1}\right)^2+\left(\frac{\partial z}{\partial q_1}\right)^2}\\ H_2&=\sqrt{\left(\frac{\partial x}{\partial q_2}\right)^2+\left(\frac{\partial y}{\partial q_2}\right)^2+\left(\frac{\partial z}{\partial q_2}\right)^2}\\ H_3&=\sqrt{\left(\frac{\partial x}{\partial q_3}\right)^2+\left(\frac{\partial y}{\partial q_3}\right)^2+\left(\frac{\partial z}{\partial q_3}\right)^2}\end{aligned}\right\} \tag{6.34}$$

we can rewrite the formulas for dl_1, dl_2 and dl_3 as follows:

$$dl_1=H_1\,dq_1, \quad dl_2=H_2\,dq_2, \quad dl_3=H_3\,dq_3 \tag{6.35}$$

The quantities H_1, H_2 and H_3 are referred to as **scale factors** (*Lamé's* coefficients*) associated with the curvilinear coordinates q_1, q_2, q_3. A coordinate curve along which only one of the parameters q_1, q_2 or q_3 varies can be thought of as a curve on which the scale of the corresponding parameter q_i ($i = 1, 2$ or 3) is marked. The multiplication by the factors H_1, H_2 and H_3 transforms infinitesimal increments of the parameters q_1, q_2 and q_3 into the corresponding increments of the parameters l_1, l_2 and l_3 which are the arc lengths of the coordinate curves.

By the hypothesis, the coordinate system being orthogonal, the area $d\sigma_1$ of the face $MM_2N_1M_3$ of the parallelepiped is equal to the product of dl_2 by dl_3, i.e.

$$d\sigma_1 = H_2 H_3 \, dq_2 \, dq_3$$

and similarly the areas $d\sigma_2$ and $d\sigma_3$ of the other two faces are expressed as

$$d\sigma_2 = H_3 H_1 \, dq_3 \, dq_1 \quad \text{and} \quad d\sigma_3 = H_1 H_2 \, dq_1 \, dq_2 \tag{6.36}$$

Finally, the volume of the infinitesimal parallelepiped is equal to

$$dv = dl_1 \, dl_2 \, dl_3 = H_1 H_2 H_3 \, dq_1 \, dq_2 \, dq_3 \tag{6.37}$$

Let us introduce, at each point M, the orthonormal basis (i.e. one whose base vectors are mutually orthogonal and have unit

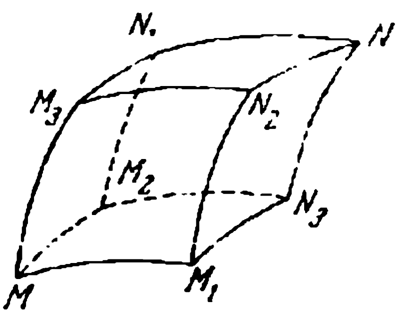


Fig. 6.13

lengths) determined by the three unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ tangent to the coordinate curves passing through the point M. It should be noted that, unlike a Cartesian coordinate system specified by three constant (both in direction and magnitude) unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ the coordinate system determined by the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ may vary from point to point in the general case, that is the vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ are functions of the coordinates q_1, q_2 and q_3. But this does not prevent us from resolving an arbitrary vector given at any point M as a linear combination of the corresponding unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$. Hence, every vector field can be (locally) resolved

* Lamé, Gabriel (1795-1870), a French scientist

into components along the directions of the vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ at each point M.

3. Cylindrical and Spherical Coordinates. Let us compute the scale factors for the cylindrical and spherical coordinate systems which are the most important special cases of curvilinear orthogonal coordinates.

Cylindrical coordinates r, φ, z are connected with Cartesian coordinates x, y, z by the formulas

$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad z = z$$

Applying formulas (6.34) we obtain

$$\left.\begin{aligned} H_1 &= \sqrt{\left(\frac{\partial x}{\partial r}\right)^2 + \left(\frac{\partial y}{\partial r}\right)^2 + \left(\frac{\partial z}{\partial r}\right)^2} = \sqrt{\cos^2\varphi + \sin^2\varphi} = 1 \\ H_2 &= \sqrt{\left(\frac{\partial x}{\partial \varphi}\right)^2 + \left(\frac{\partial y}{\partial \varphi}\right)^2 + \left(\frac{\partial z}{\partial \varphi}\right)^2} = \\ &= \sqrt{r^2\sin^2\varphi + r^2\cos^2\varphi} = r \\ H_3 &= \sqrt{\left(\frac{\partial x}{\partial z}\right)^2 + \left(\frac{\partial y}{\partial z}\right)^2 + \left(\frac{\partial z}{\partial z}\right)^2} = \sqrt{1} = 1 \end{aligned}\right\} \quad (6.38)$$

Taking into account the geometric significance of the scale factors H_1, H_2, H_3 we can directly derive formulas (6.38) without differentiation. Indeed, consider an infinitesimal parallelepiped bounded by the three pairs of the coordinate surfaces corresponding to the values r and $r + dr$, φ and $\varphi + d\varphi$, z and $z + dz$ of the cylindrical

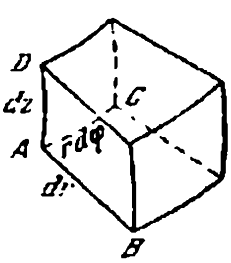


Fig. 6.14

coordinates r, φ, z (Fig. 6.14). The lengths dl_1, dl_2 and dl_3 of the edges AB, AC and AD of the parallelepiped are respectively equal to dr, $r\,d\varphi$ and dz, which immediately implies formulas (6.38).

Similarly, for the spherical coordinates determined by the relations

$$x = \rho\cos\varphi\sin\theta, \quad y = \rho\sin\varphi\sin\theta, \quad z = \rho\cos\theta$$

we find, by differentiating the formulas, the equalities

$$H_1 = 1, \quad H_2 = \rho, \quad H_3 = \rho\sin\theta \qquad (6.39)$$

This result can again be obtained directly from Fig. 6.15 since the lengths dl_1, dl_2 and dl_3 of the edges AB, AC and AD of the parallelepiped bounded by the coordinate surfaces specified by the values ρ and $\rho + d\rho$, φ and $\varphi + d\varphi$, θ and $\theta + d\theta$ of the spherical coor-

dinates ρ, φ, θ are respectively equal to

$$d\rho, \ \rho \, d\theta \quad \text{and} \quad \rho \sin \theta \, d\varphi$$

whence formulas (6.39) immediately follow.

4. Gradient. Let us find the expression of the gradient in curvilinear orthogonal coordinates. As is known, the projection of the gradient of a scalar function $U = U(q_1, q_2, q_3)$ on an arbitrary

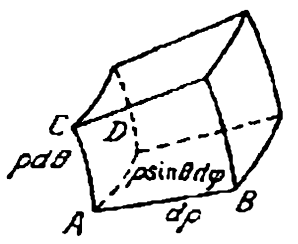


Fig. 6.15

axis coincides with the derivative of U along the direction of the axis. Consequently, to compute the components of the vector grad U with respect to the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ we must find the derivatives of U along the directions specified by these vectors. Let ΔU be the difference between the values of the function U at the points M_1 and M. Then

$$(\operatorname{grad} U, \mathbf{e}_1) = \lim_{dl_1 \to 0} \frac{\Delta U}{dl_1} = \lim_{dq_1 \to 0} \frac{\Delta U}{H \, dq_1} = \frac{1}{H_1} \frac{\partial U}{\partial q_1}$$

Similarly, the other two components of the gradient are equal to

$$\frac{1}{H_2} \frac{\partial U}{\partial q_2} \quad \text{and} \quad \frac{1}{H_3} \frac{\partial U}{\partial q_3}$$

Hence, we finally obtain

$$\operatorname{grad} U = \frac{1}{H_1} \frac{\partial U}{\partial q_1} \mathbf{e}_1 + \frac{1}{H_2} \frac{\partial U}{\partial q_2} \mathbf{e}_2 + \frac{1}{H_3} \frac{\partial U}{\partial q_3} \mathbf{e}_3 \qquad (6.40)$$

5. Divergence. We now proceed to calculate the divergence of a vector field $\mathbf{A}$ in coordinates q_1, q_2, q_3. The expression of div $\mathbf{A}$ at a point M was defined in § 2 by the formula

$$\operatorname{div} \mathbf{A} = \lim_{\Omega \to M} \frac{1}{V(\Omega)} \iint_{\Sigma} A_n \, d\sigma$$

Consequently, we can compute div $\mathbf{A}$ at an arbitrary point M by evaluating the ratio of the flux of the vector $\mathbf{A}$ through the surface of an infinitesimal parallelepiped shown in Fig. 6.13 to the volume dv of the parallelepiped. Let A_1, A_2 and A_3 be the components (coordinates, or projections) of the vector $\mathbf{A}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, i.e. let

$$\mathbf{A} = A_1 \mathbf{e}_1 + A_2 \mathbf{e}_2 + A_3 \mathbf{e}_3$$

To begin with, we find the flux of the vector $\mathbf{A}$ through the two faces perpendicular to the edge MM_1. The outer normal to the face

$MM_2N_1M_3$ coincides with the vector $-\mathbf{e}_1$ because the vector $\mathbf{e}_1$ goes in the direction of the increase of the coordinate q_1 whereas the outer normal to this face is in the opposite direction. Hence, the flux of the vector $\mathbf{A}$ through this face is given, to within infinitesimals of order higher than the first relative to the quantity $dq_2\,dq_3$, by the expression

$$(\mathbf{A},\ -\mathbf{e}_1)\,d\sigma_1 = -A_1H_2H_3\,dq_2\,dq_3 \tag{6.41}$$

where the values of A_1, H_2 and H_3 are taken at the point (q_1, q_2, q_3).

The face $M_1N_3NN_2$, opposite to the face $MM_2N_1M_3$, differs from the latter in the value of the coordinate q_1 which is now equal to $q_1 + dq_1$ instead of q_1. Consequently, the quantity $A_1H_2H_3$ takes, on the face $M_1N_3NN_2$, the value which differs from that on the face $MM_2N_1M_3$ by the increment

$$\frac{\partial}{\partial q_1}(A_1H_2H_3)\,dq_1$$

Furthermore, the outer normal to the face $M_1N_3NN_2$ is in the direction of the vector $\mathbf{e}_1$. Therefore the flux of the vector $\mathbf{A}$ through the face $M_1N_3NN_2$ is equal to

$$\left[AH_2H_3 + \frac{\partial}{\partial q_1}(A_1H_2H_3)\,dq_1\right]dq_2\,dq_3 \tag{6.42}$$

Adding together expressions (6.41) and (6.42) we find that the flux of the vector field $\mathbf{A}$ through the two opposite faces $MM_2N_1M_3$ and $M_1N_3NN_2$ is equal to

$$\frac{\partial}{\partial q_1}(A_1H_2H_3)\,dq_1\,dq_2\,dq_3$$

the last expression being correct to within infinitesimals of higher order relative to $dv = H_1H_2H_3\,dq_1\,dq_2\,dq_3$.

Similarly, taking the other two pairs of opposite faces we find the following expressions of the flux:

$$\frac{\partial(A_2H_3H_1)}{\partial q_2}\,dq_1\,dq_2\,dq_3 \quad \text{and} \quad \frac{\partial(A_3H_1H_2)}{\partial q_3}\,dq_1\,dq_2\,dq_3$$

Adding up the three quantities thus found, dividing the sum by dv and passing to the limit as $\Omega \to M$ we finally obtain the formula

$$\operatorname{div}\mathbf{A} = \frac{1}{H_1H_2H_3}\left[\frac{\partial(A_1H_2H_3)}{\partial q_1} + \frac{\partial(A_2H_3H_1)}{\partial q_2} + \frac{\partial(A_3H_1H_2)}{\partial q_3}\right] \tag{6.43}$$

6. **Rotation.** As was shown in § 4, the projection $(\operatorname{rot}\mathbf{A})_n$ of the rotation of a vector field $\mathbf{A}$ on the axis specified by an arbitrary fixed vector $\mathbf{n}$ at a point M is expressed by the formula

$$(\operatorname{rot}\mathbf{A})_n = \lim_{\Sigma\to M}\frac{\int_L A_\tau\,dl}{\sigma}$$

where Σ is a surface perpendicular to the vector $\mathbf{n}$ at the point M, σ being the area and L the boundary of Σ. Thus, we can find the projection of rot $\mathbf{A}$ on the axis of the vector $\mathbf{e}_1$ by computing the ratio of the circulation of $\mathbf{A}$ over the contour $MM_2N_1M_3M$ (shown in Fig. 6.13) to the area σ_1. Let us present the circulation as the sum of the four terms corresponding to the line segments MM_2, M_2N_1, N_1M_3, M_3M and compute each summand separately. We first take MM_2. The projection of the vector $\mathbf{A}$ on the line segment MM_2 is equal to A_2 and consequently the line integral of the projection along the path MM_2 is given, to within infinitesimals of order of smallness higher than the first relative to the length of MM_2, by the expression

$$A_2\,dl_2 = A_2H_2\,dq_2 \tag{6.44}$$

where the values of the quantities A_2 and H_2 are taken at the point (q_1, q_2, q_3). The integral over N_1M_3 differs from the above expression in the third coordinate, which assumes on N_1M_3 the value $q_3 + dq_3$ (instead of the value q_3 taken on MM_2), and in the direction of integration since the direction of the line segment N_1M_3 is opposite to that of the vector $\mathbf{e}_2$. Therefore the integral of the projection of $\mathbf{A}$ on N_1M_3 over the path N_1M_3 is equal to

$$-\left[A_2H_2 + \frac{\partial}{\partial q_3}(A_2H_2)\,dq_3\right]dq_2 \tag{6.45}$$

Similarly, for the integrals taken along M_2N_1 and M_3M we obtain the corresponding expressions

$$\left[A_3H_3 + \frac{\partial}{\partial q_2}(A_3H_3)\,dq_2\right]dq_3 \tag{6.46}$$

and

$$-A_3H_3\,dq_3 \tag{6.47}$$

Adding together expressions (6.44), (6.45), (6.46), and (6.47) we find that the circulation of the vector $\mathbf{A}$ over the contour $MM_2N_1M_3M$ is equal to

$$-\frac{\partial(A_2H_2)}{\partial q_3}\,dq_2\,dq_3 + \frac{\partial(A_3H_3)}{\partial q_2}\,dq_2\,dq_3$$

the last expression being correct to within infinitesimals of higher order relative to $dq_2\,dq_3$. Dividing this expression by $H_2H_3\,dq_2\,dq_3$, i.e. by the area of the face $MM_2N_1M_3$, and passing to the limit as $\Sigma \to M$ we see that the component $(\text{rot}\,\mathbf{A})_1$ of the vector rot $\mathbf{A}$ in the direction of the base vector $\mathbf{e}_1$ is equal to

$$\frac{1}{H_2H_3}\left\{\frac{\partial(A_3H_3)}{\partial q_2} - \frac{\partial(A_2H_2)}{\partial q_3}\right\} \tag{6.48$_1$}$$

The other two components are computed analogously:

$$(\text{rot}\,\mathbf{A})_2 = \frac{1}{H_3H_1}\left\{\frac{\partial(A_1H_1)}{\partial q_3} - \frac{\partial(A_3H_3)}{\partial q_1}\right\} \tag{6.48$_2$}$$

and

$$(\text{rot } \mathbf{A})_3 = \frac{1}{H_1 H_2} \left\{ \frac{\partial (A_2 H_2)}{\partial q_1} - \frac{\partial (A_1 H_1)}{\partial q_2} \right\} \qquad (6.48_3)$$

Finally we can write the resultant expression of the rotation in curvilinear orthogonal coordinates q_1, q_2, q_3:

$$\text{rot } \mathbf{A} = \frac{1}{H_2 H_3} \left\{ \frac{\partial (A_3 H_3)}{\partial q_2} - \frac{\partial (A_2 H_2)}{\partial q_3} \right\} \mathbf{e}_1 + \frac{1}{H_3 H_1} \left\{ \frac{\partial (A_1 H_1)}{\partial q_3} - \frac{\partial (A_3 H_3)}{\partial q_1} \right\} \mathbf{e}_2 + \\ + \frac{1}{H_1 H_2} \left\{ \frac{\partial (A_2 H_2)}{\partial q_1} - \frac{\partial (A_1 H_1)}{\partial q_2} \right\} \mathbf{e}_3$$

7. **Laplace's Operator.** Based on the expressions of grad U and div $\mathbf{A}$, we can write the following formula for Laplace's operator in curvilinear orthogonal coordinates q_1, q_2, q_3:

$$\Delta U = \text{div grad } U = \frac{1}{H_1 H_2 H_3} \left\{ \frac{\partial}{\partial q_1} \left(\frac{H_2 H_3}{H_1} \frac{\partial U}{\partial q_1} \right) + \right. \\ \left. + \frac{\partial}{\partial q_2} \left(\frac{H_3 H_1}{H_2} \frac{\partial U}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left(\frac{H_1 H_2}{H_3} \frac{\partial U}{\partial q_3} \right) \right\} \qquad (6.49)$$

To memorize this formula note that the scale factors H_1, H_2 and H_3 entering into the denominators of the expressions written in front of the derivatives $\frac{\partial U}{\partial q_i}$ ($i = 1, 2, 3$) are due to the gradient, the factors $H_2 H_3$, $H_3 H_1$ and $H_1 H_2$ in the numerators are generated by the expressions of the areas of the faces (which have been taken when computing the fluxes) and the factor $\frac{1}{H_1 H_2 H_3}$ has appeared when we have divided the resultant flux across the faces by the volume of the infinitesimal parallelepiped.

8. **Basic Field Operations in Cylindrical and Spherical Coordinates.** In Sec. 3 we found the scale factors for the cylindrical and spherical coordinates. To write down the formulas expressing the gradient, divergence, rotation and Laplacian operator in these coordinates we must only substitute the corresponding scale factors into the general formulas obtained in Secs. 4-7. This yields the following results:

(a) In cylindrical coordinates we have

$$\text{grad } U = \frac{\partial U}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial U}{\partial \varphi} \mathbf{e}_\varphi + \frac{\partial U}{\partial z} \mathbf{e}_z,$$

$$\text{div } \mathbf{A} = \frac{1}{r} \frac{\partial (r A_r)}{\partial r} + \frac{1}{r} \frac{\partial A_\varphi}{\partial \varphi} + \frac{\partial A_z}{\partial z},$$

$$\text{rot } \mathbf{A} = \left(\frac{1}{r} \frac{\partial A_z}{\partial \varphi} - \frac{\partial A_\varphi}{\partial z} \right) \mathbf{e}_r + \left(\frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right) \mathbf{e}_\varphi + \\ + \left(\frac{1}{r} \frac{\partial (r A_\varphi)}{\partial r} - \frac{1}{r} \frac{\partial A_r}{\partial \varphi} \right) \mathbf{e}_z,$$

$$\Delta U = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial U}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 U}{\partial \varphi^2} + \frac{\partial^2 U}{\partial z^2}$$

where $\mathbf{e}_r$, $\mathbf{e}_\varphi$ and $\mathbf{e}_z$ designate, respectively, the vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ corresponding to the cylindrical coordinates r, φ, z and A_r, A_φ and A_z are the projections of $\mathbf{A}$ on $\mathbf{e}_r$, $\mathbf{e}_\varphi$ and $\mathbf{e}_z$.

(b) In spherical coordinates we obtain

$$\operatorname{grad} U = \frac{\partial U}{\partial \rho}\mathbf{e}_\rho + \frac{1}{\rho}\frac{\partial U}{\partial \theta}\mathbf{e}_\theta + \frac{1}{\rho\sin\theta}\frac{\partial U}{\partial \varphi}\mathbf{e}_\varphi,$$

$$\operatorname{div}\mathbf{A} = \frac{1}{\rho^2}\frac{\partial(\rho^2 A_\rho)}{\partial\rho} + \frac{1}{\rho\sin\theta}\frac{\partial(\sin\theta\, A_\theta)}{\partial\theta} + \frac{1}{\rho\sin\theta}\frac{\partial A_\varphi}{\partial\varphi},$$

$$\operatorname{rot}\mathbf{A} = \frac{1}{\rho\sin\theta}\left(\frac{\partial(A_\varphi\sin\theta)}{\partial\theta} - \frac{\partial A_\theta}{\partial\varphi}\right)\mathbf{e}_\rho + \left(\frac{1}{\rho\sin\theta}\frac{\partial A_\rho}{\partial\varphi} - \frac{1}{\rho}\frac{\partial(\rho A_\varphi)}{\partial\rho}\right)\mathbf{e}_\theta + \left(\frac{1}{\rho}\frac{\partial(\rho A_\theta)}{\partial\rho} - \frac{1}{\rho}\frac{\partial A_\rho}{\partial\theta}\right)\mathbf{e}_\varphi,$$

$$\Delta U = \frac{1}{\rho^2}\frac{\partial}{\partial\rho}\left(\rho^2\frac{\partial U}{\partial\rho}\right) + \frac{1}{\rho^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial U}{\partial\theta}\right) + \frac{1}{\rho^2\sin^2\theta}\frac{\partial^2 U}{\partial\varphi^2}$$

where $\mathbf{e}_\rho$, $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$ denote the base vectors of the spherical coordinate system and A_ρ, A_θ, A_φ are the corresponding projections of $\mathbf{A}$.

In some problems related to the Laplacian operator, besides the whole expression of ΔU in spherical coordinates, we encounter its part of the form

$$\frac{1}{\rho^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial U}{\partial\theta}\right) + \frac{1}{\rho^2\sin^2\theta}\frac{\partial^2 U}{\partial\varphi^2}$$

involving only the derivatives with respect to the "angular" variables θ and φ.

Note. In § 7 we systematically used such terms as "an infinitesimal parallelepiped", "an element of volume" and the like. It is clear that here, as in many other similar cases, the real sense of this terminology is that we first take the corresponding geometrical and physical objects of finite sizes and then perform the passage to the limit making the sizes tend to zero. We believe that the reader can easily perform all the passages to the limit we are speaking about and which have not been written in full in the foregoing sections.

§ 8. VARIABLE FIELDS IN CONTINUOUS MEDIA

So far, in studying various fields we investigated the dependence of the corresponding quantities (scalar or vector) on the spatial coordinates. Here we are going to consider some questions related to the dependence of fields on time.

1. Partial and Total Time Derivatives. Let us consider a fluid motion whose velocity depends not only on the coordinates at each point but also on time. Suppose there is a variable quantity φ related to a fluid flow, for instance, temperature, pressure etc. The variation

of the quantity can be studied by means of two different approaches: we can investigate its variation at a given point, i.e. for some fixed values of x, y and z, or consider the values of the quantity connected with a certain fluid particle whose coordinates are time-variable. For example, if we deal with the temperature of a flow of fluid we can measure it with a thermometer fixed at a certain point or with one going downstream with the fluid.

The variation in time of a quantity $\varphi(M, t)$ at a fixed point M is characterized by the *partial derivative* (also termed *explicit partial derivative*)

$$\frac{\partial\varphi}{\partial t}=\lim_{\Delta t\to 0}\frac{\varphi(M,\ t+\Delta t)-\varphi(M,\ t)}{\Delta t} \tag{6.50}$$

expressing the local rate of change of φ, the coordinates x, y and z of the point M being regarded as constant parameters when computing (6.50).

The time-variation of the quantity $\varphi(M, t)$ connected with a certain particle of fluid is specified by the *total partial derivative* (*particle derivative*) of $\varphi(M, t)$ with respect to t which is defined as follows.

Let a particle be at a point M at moment t and at a point M' at moment $t+\Delta t$. **The total partial derivative of φ with respect to t** is the limit

$$\frac{d\varphi}{dt}=\lim_{\Delta t\to 0}\frac{\varphi(M',\ t+\Delta t)-\varphi(M,\ t)}{\Delta t} \tag{6.51}$$

To establish the connection between the partial and total time derivatives we note that when computing the total derivative we must consider the coordinates x, y and z of the point M to be functions of t whose derivatives with respect to t are the components of the velocity of the fluid flow at the point:

$$\frac{dx}{dt}=v_x,\qquad \frac{dy}{dt}=v_y,\qquad \frac{dz}{dt}=v_z$$

Therefore we must differentiate $\varphi=\varphi(x, y, z, t)$ as a composite function of t, which results in

$$\frac{d\varphi}{dt}=\frac{\partial\varphi}{\partial t}+\frac{\partial\varphi}{\partial x}v_x+\frac{\partial\varphi}{\partial y}v_y+\frac{\partial\varphi}{\partial z}v_z$$

that is

$$\frac{d\varphi}{dt}=\frac{\partial\varphi}{\partial t}+(\mathbf{v},\ \mathrm{grad}\ \varphi) \tag{6.52}$$

The notions of a partial and a total derivative with respect to time can be analogously introduced for an arbitrary vector quantity $\mathbf{A}(M, t)$ connected with a moving fluid. These derivatives are determined by the formulas

$$\frac{\partial\mathbf{A}}{\partial t}=\lim_{\Delta t\to 0}\frac{\mathbf{A}(M,\ t+\Delta t)-\mathbf{A}(M,\ t)}{\Delta t} \tag{6.53}$$

and

$$\frac{d\mathbf{A}}{dt} = \lim_{\Delta t \to 0} \frac{\mathbf{A}(M, t+\Delta t) - \mathbf{A}(M, t)}{\Delta t} \tag{6.54}$$

which are similar to formulas (6.50) and (6.51). The relationship between the derivatives is given by the formula

$$\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + \frac{\partial \mathbf{A}}{\partial x} v_x + \frac{\partial \mathbf{A}}{\partial y} v_y + \frac{\partial \mathbf{A}}{\partial z} v_z \tag{6.55}$$

which is obtained by differentiating $\mathbf{A}(x, y, z, t)$ as a composite function of t. Equalities (6.52) and (6.55) can be conveniently rewritten in the form

$$\frac{d\varphi}{dt} = \frac{\partial \varphi}{\partial t} + (\mathbf{v}, \nabla)\varphi \tag{6.56}$$

and

$$\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v}, \nabla)\mathbf{A} \tag{6.57}$$

where the expression $(\mathbf{v}, \nabla)$ is understood as the operator

$$v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} + v_z \frac{\partial}{\partial z}$$

i.e. as the (formal) scalar product of the velocity vector $\mathbf{v}$ by the symbolic vector ∇.

The expressions $(\mathbf{v}, \nabla)\varphi$ and $(\mathbf{v}, \nabla)\mathbf{A}$ entering into formulas (6.56) and (6.57) are *convective terms* because they are connected with convection of the particles. They only appear when we deal with a moving medium.

As an example, let us consider the acceleration of the particles of a moving liquid. It can be obtained as the total (particle) derivative of the velocity with respect to time. Taking advantage of formula (6.56) we derive

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}, \nabla)\mathbf{v} \tag{6.58}$$

or, in the coordinate notation,

$$\frac{dv_x}{dt} = \frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} + v_z \frac{\partial v_x}{\partial z},$$

$$\frac{dv_y}{dt} = \frac{\partial v_y}{\partial t} + v_x \frac{\partial v_y}{\partial x} + v_y \frac{\partial v_y}{\partial y} + v_z \frac{\partial v_y}{\partial z},$$

$$\frac{dv_z}{dt} = \frac{\partial v_z}{\partial t} + v_x \frac{\partial v_z}{\partial x} + v_y \frac{\partial v_z}{\partial y} + v_z \frac{\partial v_z}{\partial z}$$

2. Eulerian Equations of Motion of Ideal Liquid. Let us apply the notions of partial and total time derivatives to obtaining the

equations of motion of an ideal liquid. We shall consider the so-called **Eulerian specification of fluid motion** playing an important role in hydrodynamics.

Take a volume Ω bounded by a surface Σ lying inside a moving liquid. An element of area $d\sigma$ of the surface is acted upon by a force produced by the pressure. This force is directed along the normal $\mathbf{n}$ to $d\sigma$* and is equal to $-p\mathbf{n}\,d\sigma$. (Here $\mathbf{n}$ is the unit vector in the direction of the outer normal to Σ and p is the pressure, a scalar quantity.) The resultant force $\mathbf{F}$ acting upon the whole surface Σ is expressed in the form

$$\mathbf{F} = -\iint_{\Sigma} p\mathbf{n}\,d\sigma \tag{6.59}$$

where, as before, the integral of the vector

$$p\mathbf{n} = p\cos(\mathbf{n}, x)\,\mathbf{i} + p\cos(\mathbf{n}, y)\,\mathbf{j} + p\cos(\mathbf{n}, z)\,\mathbf{k}$$

is understood as a vector whose components are

$$\iint_{\Sigma} p\cos(\mathbf{n}, x)\,d\sigma,\quad \iint_{\Sigma} p\cos(\mathbf{n}, y)\,d\sigma \text{ and } \iint_{\Sigma} p\cos(\mathbf{n}, z)\,d\sigma \tag{6.60}$$

Surface integral (6.59) can be reduced to the corresponding volume integral over Ω by applying the Ostrogradsky theorem to each component (6.60) of the integral (which, as has been mentioned, is a vector). This yields

$$-\iint_{\Sigma} p\mathbf{n}\,d\sigma = -\mathbf{i}\iiint_{\Omega} \frac{\partial p}{\partial x}\,d\omega - \mathbf{j}\iiint_{\Omega} \frac{\partial p}{\partial y}\,d\omega -$$
$$-\mathbf{k}\iiint_{\Omega} \frac{dp}{dz}\,d\omega = -\iiint_{\Omega} \operatorname{grad} p\,d\omega$$

and consequently an element of volume $d\omega$ of the liquid is subjected to the force

$$-\operatorname{grad} p\,d\omega$$

On the other hand, if $\rho(M, t)$ is the mass density of the liquid at an arbitrary point M at moment t and $\mathbf{w}$ is the acceleration of a particle passing through the point M, the quantity $\mathbf{w}\rho(M, t)\,d\omega$ is the product of the mass contained in the volume $d\omega$ by the acceleration and hence, by Newton's second law, we obtain the equality

$$\mathbf{w}\,\rho\,d\omega = -\operatorname{grad} p\,d\omega$$

---

* We suppose that the liquid is *ideal* which means that its *viscosity factor* is considered to be equal to zero. Therefore the force acting upon an infinitesimal area placed inside the liquid is only produced by pressure and is directed perpendicularly to the area.

i.e.

$$\mathbf{w}\rho = -\operatorname{grad} p \tag{6.61}$$

This very equation yields an Eulerian description of motion of an ideal liquid.* Here $\mathbf{w}$ is understood as the acceleration of a particle of liquid, that is as the total derivative $\mathbf{w} = \frac{d\mathbf{v}}{dt}$ of the velocity $\mathbf{v}$ with respect to time t. Using the expressions for the components of the acceleration found in Sec. 1 we can rewrite equation (6.61) in coordinate form:

$$\left(\frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} + v_z \frac{\partial v_x}{\partial z}\right)\rho = -\frac{\partial p}{\partial x},$$

$$\left(\frac{\partial v_y}{\partial t} + v_x \frac{\partial v_y}{\partial x} + v_y \frac{\partial v_y}{\partial y} + v_z \frac{\partial v_y}{\partial z}\right)\rho = -\frac{\partial p}{\partial y},$$

$$\left(\frac{\partial v_z}{\partial t} + v_x \frac{\partial v_z}{\partial x} + v_y \frac{\partial v_z}{\partial y} + v_z \frac{\partial v_z}{\partial z}\right)\rho = -\frac{\partial p}{\partial z}$$

3. Derivative with Respect to Time of an Integral Over a Fluid Volume. Let us consider a volume Ω of a moving continuous medium. We shall refer to Ω as a *fluid volume* if it constantly consists of the same particles of the fluid. A fluid volume thus may move and change its shape in the process of motion. Consider the integral

$$J = \iiint\limits_{\Omega} \varphi \, d\omega \tag{6.62}$$

of a scalar function $\varphi(M, t)$ taken over such a volume. We shall compute the derivative of this integral with respect to time.

In performing the computation we must take into account that the variation of integral (6.62) in time is specified by two factors, namely by the variation of the integrand connected with the increment of time t and by the change of the shape and position of the spatial domain Ω the integral is taken over.**

If the volume Ω were invariable the function φ would gain the increment $\frac{\partial \varphi}{\partial t} dt$ in time dt and integral (6.62) would acquire the

---

* The minus sign on the right-hand side of equation (6.61) has an obvious physical significance, namely it indicates that the acceleration of every particle of liquid is in the direction of decrease of the pressure, i.e. opposite to the gradient of p.

** Expression (6.62) is the so-called *integral dependent on a parameter*, the parameter being t. (In (6.62) both the integrand and the domain of integration depend on t.) The fundamentals of the theory of integrals involving a parameter will be presented in Chapter 10. Here, without resorting to the general theory, we shall only consider the question of finding the derivative of integral (6.62) with respect to time, which is important for various physical applications.

increment

$$dt \iiint\limits_{\Omega} \frac{\partial \varphi}{\partial t}\, d\omega$$

Now consider the increment of integral (6.62) produced by the variation of the domain Ω. Let Σ designate the surface bounding the volume Ω at moment t. The variation of the volume Ω during time interval from t to $t + dt$ is obviously due to the fact that some particles of the fluid flow into or out of the volume Ω through the surface Σ. The volume of the fluid flowing out through an element $d\sigma$ of the surface Σ during time dt is equal to $v_n\, dt\, d\sigma$ where v_n is the projection of the velocity of the fluid on the outer normal to $d\sigma$. This variation of the volume gives, to integral (6.62), the increment

$$\varphi v_n\, dt\, d\sigma$$

The resultant increment of integral (6.62) due to the variation of the volume Ω in time dt is equal to

$$dt \iint\limits_{\Sigma} \varphi v_n\, d\sigma$$

Thus, the total increment of integral (6.62) during time period dt is equal to

$$dJ = dt \iiint\limits_{\Omega} \frac{\partial \varphi}{\partial t}\, d\omega + dt \iint\limits_{\Sigma} \varphi v_n\, d\sigma$$

and, consequently, we have

$$\frac{dJ}{dt} = \iiint\limits_{\Omega} \frac{\partial \varphi}{\partial t}\, d\omega + \iint\limits_{\Sigma} \varphi v_n\, d\sigma \tag{6.63}$$

Transforming the second summand on the right-hand side by the Ostrogradsky theorem we obtain

$$\frac{d}{dt} \iiint\limits_{\Omega} \varphi\, d\omega = \iiint\limits_{\Omega} \left[\frac{\partial \varphi}{\partial t} + \operatorname{div}(\varphi \mathbf{v}) \right] d\omega \tag{6.64}$$

Finally, taking advantage of the equality

$$\operatorname{div}(\varphi \mathbf{v}) = \varphi \operatorname{div} \mathbf{v} + (\mathbf{v}, \operatorname{grad} \varphi)$$

(see formula (6.29)) and applying expression (6.52) for the total derivative we receive the resulting formula

$$\frac{d}{dt} \iiint\limits_{\Omega} \varphi\, d\omega = \iiint\limits_{\Omega} \left(\frac{d\varphi}{dt} + \varphi \operatorname{div} \mathbf{v} \right) d\omega \tag{6.65}$$

In particular, if $\operatorname{div} \mathbf{v} = 0$, i.e. if we consider a motion of an incompressible liquid without sources and sinks, formula (6.65)

takes a simpler form

$$\frac{d}{dt}\iiint\limits_{\Omega}\varphi\, d\omega=\iiint\limits_{\Omega}\frac{d\varphi}{dt}\, d\omega$$

Note. The problem of differentiating an integral taken over a fluid volume is analogous to the following one-dimensional problem (which we shall again deal with in Chapter 10): given an integral

$$J(t)\equiv\int\limits_{a(t)}^{b(t)}\varphi(x,\, t)\, dx$$

it is required to find the derivative of $J(t)$ with respect to t. Regarding $J(t)$ as a composite function of t (dependent on t because both the integrand and the limits of integration $a(t)$ and $b(t)$ involve t) we easily find that

$$J'(t)=\int\limits_{a(t)}^{b(t)}\frac{\partial\varphi}{\partial t}\, dx+\varphi(b(t),\, t)\, b'(t)-\varphi(a(t),\, t)\, a'(t)$$

Here again $J'(t)$ is a sum of two terms $\left(\int\limits_{a(t)}^{b(t)}\frac{\partial\varphi}{\partial t}\, dx\right.$ and $\left.[\varphi(b,\, t)\, b'-\varphi(a,\, t)\, a']\right)$, the former being due to the variation of the integrand and the latter to the change of the interval of integration.

We have studied the integral of a scalar function over a fluid volume. Similar techniques can be applied to studying an integral of a vector function $\mathbf{A}(M,\, t)$. The same arguments yield the following expression for the derivative of such an integral with respect to t

$$\frac{d}{dt}\iiint\limits_{\Omega}\mathbf{A}\, d\omega=\iiint\limits_{\Omega}\left[\frac{d\mathbf{A}}{dt}+\mathbf{A}\operatorname{div}\mathbf{v}\right]d\omega \qquad (6.66)$$

The above discussion concerns integration over a volume of fluid. But in hydrodynamics and some other divisions of physics we also deal with surfaces and curves formed of particles of fluid which change their position in space and their shape in accordance with the motion of the particles. Surface and line integrals of functions over such fluid surfaces and curves are again expressions whose dependence on time is specified by two factors, i.e. by the variation, in time, of the domains of integration and by that of the integrands. Using arguments similar to those applied to investigating integrals (6.62) and (6.66) we can easily establish the corresponding formulas for differentiating these surface and line integrals with respect to time.

4. Application to Deriving Equation of Continuity. Formula (6.63) immediately implies the equation of continuity obtained in § 3, Sec. 5. Indeed, let $\rho(M, t)$ be the density of a moving liquid (which may be compressible in the general case). The mass of the liquid contained in a volume Ω is equal to

$$\iiint\limits_{\Omega} \rho\, d\omega$$

If Ω is a fluid volume which is thought of as consisting of the same particles of the fluid in the process of motion, the mass it contains remains constant. Consequently, by formula (6.64), we have

$$\frac{d}{dt}\iiint\limits_{\Omega} \rho\, d\omega = \iiint\limits_{\Omega} \left[\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v})\right] d\omega = 0$$

The volume Ω being taken quite arbitrarily, we thus obtain the relation

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0$$

which is the equation of continuity.

7 Tensors

The investigation of phenomena in natural and engineering sciences involves various mathematical quantities. The distinction between these quantities lies in their analytic expressions and the laws of transformation of the expressions when the coordinate system is changed.

From the point of view of mathematics, the simplest physical quantities are *scalars* (such as the mass of a body, the volume of a body, the length of a vector etc.) which are invariant with respect to the transformations of coordinates. Every scalar quantity is characterized, in any coordinate system, by a single numerical value independent of the choice of the coordinate system.

The vector quantities, e.g. velocity, acceleration, force etc., are of a more complicated mathematical nature. A vector quantity is specified, in the three-dimensional space, by a triple of numbers with respect to each coordinate basis, namely by the three projections of the vector on the coordinate axes, which are also referred to as the *coordinates* (or *components*) of the vector. When a coordinate basis is replaced by another one the components of a vector are transformed according to a special law.

Still more complicated quantities, from the point of view of the laws of their transformation, are the so-called *tensors* which, in certain particular cases, are analogous to linear operators applied to vectors (the notion of a linear operator is discussed in Sec. 1 of § 2 in the present chapter).

As an example of such a quantity, let us consider the so-called *conductivity tensor* characterizing electric properties of an anisotropic conductor. In an isotropic medium the current density vector $\mathbf{j}$ and the electric intensity vector $\mathbf{E}$ are collinear, that is connected by a relation

$$\mathbf{j} = \sigma \mathbf{E} \tag{7.1}$$

where $\sigma > 0$ is a scalar factor known as the *specific conductivity* of the medium. But in the general case of an anisotropic body the vectors $\mathbf{j}$ and $\mathbf{E}$ are no longer collinear and therefore the factor σ should be regarded as a linear operator transforming the vector $\mathbf{E}$ into the vector $\mathbf{j}$. This operator is referred to as the conductivity tensor.

If we choose an arbitrary fixed basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in space and resolve the vectors $\mathbf{j}$ and $\mathbf{E}$ with respect to the basis we can write

$$\begin{aligned} \mathbf{j} &= j_1\mathbf{e}_1 + j_2\mathbf{e}_2 + j_3\mathbf{e}_3 \\ \mathbf{E} &= E_1\mathbf{e}_1 + E_2\mathbf{e}_2 + E_3\mathbf{e}_3 \end{aligned} \tag{7.2}$$

and then replace relation (7.1) by the equivalent system of three scalar equalities

$$j_k = \sum_{i=1}^{3} \sigma_{ki} E_i, \qquad k = 1, 2, 3 \tag{7.1'}$$

Thus, the conductivity tensor σ is specified, in every basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, by the nine numbers σ_{ki}, k, $i = 1, 2, 3$, termed the *components* of the tensor σ relative to this basis.

The definition of the notion of tensor should obviously include the rule for transforming its components when the basis is changed.

In §§ 1-9 we shall only restrict ourselves to the orthonormal bases (i.e. the ones consisting of mutually orthogonal unit vectors) and their transformations connected with the notion of an *orthogonal affine tensor*. In § 10 we shall briefly discuss the general tensors.

§ 1. ORTHOGONAL AFFINE TENSOR

1. Transformation of Orthonormal Bases. Let us consider two arbitrary orthonormal bases $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ in a three-dimensional Euclidean space. The bases being orthonormal, we have the following relations for their scalar products:

$$(\mathbf{e}_i, \mathbf{e}_k) = \delta_{ik}, \quad (\mathbf{e}'_i, \mathbf{e}'_k) = \delta_{ik}, \quad \delta_{ik} = \begin{cases} 0 & \text{for } i \neq k \\ 1 & \text{for } i = k \end{cases} \tag{7.3}$$

(δ_{ik} is called the **Kronecker*** **delta**). We shall conditionally refer to $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ as the "old" basis and to $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ as the "new" one. Resolving the base vectors of the new basis with respect to the old basis we obtain the equalities

$$\begin{aligned} \mathbf{e}'_1 &= \alpha_{11}\mathbf{e}_1 + \alpha_{12}\mathbf{e}_2 + \alpha_{13}\mathbf{e}_3 \\ \mathbf{e}'_2 &= \alpha_{21}\mathbf{e}_1 + \alpha_{22}\mathbf{e}_2 + \alpha_{23}\mathbf{e}_3 \\ \mathbf{e}'_3 &= \alpha_{31}\mathbf{e}_1 + \alpha_{32}\mathbf{e}_2 + \alpha_{33}\mathbf{e}_3 \end{aligned} \tag{7.4}$$

or, in the contracted notation,

$$\mathbf{e}'_i = \sum_{j=1}^{3} \alpha_{ij}\mathbf{e}_j, \qquad i = 1, 2, 3 \tag{7.4'}$$

---

* Kronecker, Leopold (1823-1891), a German mathematician.

The matrix

$$\|\alpha_{ij}\| = \begin{Vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{Vmatrix} \tag{7.5}$$

is called the transformation matrix from the old basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ to the new basis $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$.

Let us investigate the properties of this matrix. Multiplying scalarly the vector $\mathbf{e}'_i = \alpha_{i1}\mathbf{e}_1 + \alpha_{i2}\mathbf{e}_2 + \alpha_{i3}\mathbf{e}_3$ $(i = 1, 2, 3)$ by the vector $\mathbf{e}'_j = \alpha_{j1}\mathbf{e}_1 + \alpha_{j2}\mathbf{e}_2 + \alpha_{j3}\mathbf{e}_3$ $(j = 1, 2, 3)$ we receive

$$\alpha_{i1}\alpha_{j1} + \alpha_{i2}\alpha_{j2} + \alpha_{i3}\alpha_{j3} = \delta_{ij} = \begin{cases} 0 & \text{for } j \neq i \\ 1 & \text{for } j = i \end{cases} \tag{7.6}$$

Hence, the sum of the squares of the elements of every row of the matrix $\|\alpha_{ij}\|$ is equal to unity and the sum of the products of the corresponding elements of two different rows is equal to zero.* Next, computing the scalar products of expression (7.4′) by $\mathbf{e}_k$ $(k = 1, 2, 3)$ we obtain**

$$(\mathbf{e}'_i, \mathbf{e}_k) = \alpha_{ik}, \qquad i, k = 1, 2, 3 \tag{7.7}$$

We now proceed to find the expressions for the elements of the inverse of matrix (7.5). For this purpose we expand the base vectors of the old basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ with respect to the new basis and thus derive

$$\begin{aligned} \mathbf{e}_1 &= \beta_{11}\mathbf{e}'_1 + \beta_{12}\mathbf{e}'_2 + \beta_{13}\mathbf{e}'_3 \\ \mathbf{e}_2 &= \beta_{21}\mathbf{e}'_1 + \beta_{22}\mathbf{e}'_2 + \beta_{23}\mathbf{e}'_3 \\ \mathbf{e}_3 &= \beta_{31}\mathbf{e}'_1 + \beta_{32}\mathbf{e}'_2 + \beta_{33}\mathbf{e}'_3 \end{aligned} \tag{7.8}$$

or, in the abbreviated notation,

$$\mathbf{e}_k = \sum_{j=1}^{3} \beta_{kj}\mathbf{e}'_j \tag{7.8′}$$

The matrix

$$\|\beta_{ij}\| = \begin{Vmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{Vmatrix} \tag{7.9}$$

is apparently the inverse of matrix (7.5). Performing scalar multiplication of expression (7.8′) by $\mathbf{e}'_i$ $(i = 1, 2, 3)$ we obtain

$$(\mathbf{e}'_i, \mathbf{e}_k) = \beta_{ki}, \qquad i, k = 1, 2, 3 \tag{7.10}$$

---

* A matrix $\|\alpha_{ij}\|$ for which relations (7.6) hold is called orthogonal. Thus, a transformation matrix from one orthonormal basis to another is an orthogonal matrix.

** We obviously have $\alpha_{ik} = (\mathbf{e}'_i, \mathbf{e}_k) = \cos(\mathbf{e}'_i, \mathbf{e}_k)$ $(i, k = 1, 2, 3)$.

Comparing formulas (7.7) and (7.10) we find the following relationship between the elements of matrices (7.5) and (7.9):

$$\alpha_{ik} = \beta_{ki} \quad (i,\ k = 1,\ 2,\ 3) \tag{7.11}$$

Thus, matrix (7.9), the inverse of matrix (7.5), is the *transpose* of matrix (7.5). It follows that the columns of an orthogonal matrix possess the same properties as its rows (see formula (7.6)) since the inverse of matrix (7.5) is the transformation matrix from the orthonormal basis $\mathbf{e}_1'$, $\mathbf{e}_2'$, $\mathbf{e}_3'$ to the orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and hence it is an orthogonal matrix.

2. **Definition of Orthogonal Affine Tensor.** In constructing the formal theory of tensors it appears expedient to include the invariant scalar quantities and vectors into the class of tensors. Thus, a scalar quantity L invariant with respect to the transformations from one orthonormal basis to another is called an **orthogonal affine tensor of rank zero.**

Such quantities as temperature, mass, the length of a vector etc. are examples of orthogonal affine tensors of rank zero. On the other hand, although the projection of a vector on the ith ($i = 1,\ 2,\ 3$) coordinate axis (i.e. the axis determined by the ith base vector $\mathbf{e}_i$) is specified by a single number in each basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, it changes in a certain manner when the basis is replaced and thus it is not a tensor of rank zero.

To include the vectors into the class of tensors we define them as tensors of *rank* (*order*) *one*.

Definition 1. Let a quantity L *be determined, in every orthonormal basis, by a triple of numbers, say by numbers L_1, L_2, L_3 in a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, by numbers L_1', L_2', L_3' in a basis $\mathbf{e}_1'$, $\mathbf{e}_2'$, $\mathbf{e}_3'$ etc. Suppose that under the transformation from an arbitrary orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ to any other orthonormal basis $\mathbf{e}_1'$, $\mathbf{e}_2'$, $\mathbf{e}_3'$ these numbers undergo the transformation determined by the formulas*

$$L_i' = \sum_{k=1}^{3} \alpha_{ik} L_k \qquad (i = 1,\ 2,\ 3) \tag{7.12}$$

where $\| \alpha_{ik} \|$ is the transformation matrix from the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ to the basis $\mathbf{e}_1'$, $\mathbf{e}_2'$, $\mathbf{e}_3'$. Then the quantity L *is said to be an* ***orthogonal affine tensor of rank (order) one*** *and is denoted by the symbol (L_i), i.e.* L $\equiv (L_i)$.

The numbers L_i, $i = 1,\ 2,\ 3$, are referred to as the **components** (coordinates) of the tensor L, relative to the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, the numbers L_i', $i = 1,\ 2,\ 3$, as the components in the basis $\mathbf{e}_1'$, $\mathbf{e}_2'$, $\mathbf{e}_3'$ etc.

Let us prove that every vector is an orthogonal affine tensor of rank (order) one. Indeed, every vector $\mathbf{x}$ is determined by a triple of numbers in each orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, these numbers being

its coordinates relative to the basis, i.e. its projections on the corresponding coordinate axes. Besides, when the orthonormal basis is changed the coordinates of the vector $\mathbf{x}$ are transformed according to formulas of type (7.12). For, if we resolve $\mathbf{x}$ with respect to two bases $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ we obtain

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = x'_1\mathbf{e}'_1 + x'_2\mathbf{e}'_2 + x'_3\mathbf{e}'_3 \tag{7.13}$$

Multiplying scalarly equality (7.13) by $\mathbf{e}'_i$ ($i = 1, 2, 3$) we find, with the help of formulas (7.3) and (7.7), that

$$x'_i = \alpha_{i1}x_1 + \alpha_{i2}x_2 + \alpha_{i3}x_3 = \sum_{k=1}^{3} \alpha_{ik}x_k, \qquad i = 1, 2, 3 \tag{7.14}$$

Formulas (7.14) being of type (7.12), we thus see that every vector $\mathbf{x}$ is in fact an orthogonal affine tensor of rank one.

Note 1. It is obvious that, conversely, every orthogonal affine tensor of rank one can be interpreted as a vector.

Note 2. Since the inverse of the matrix $\|\alpha_{ij}\|$ is its transpose, relations (7.14) imply that

$$x_i = \sum_{j=1}^{3} \alpha_{ji}x'_j, \qquad i = 1, 2, 3 \tag{7.14'}$$

We now proceed to define a tensor of second rank.

Definition 2. *Let a quantity* L *be determined by nine numbers in every orthonormal basis, these numbers being* L_{ij}, $i, j = 1, 2, 3$, *in a basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, L'_{ij}, $i, j = 1, 2, 3$, *in another basis* $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ *etc. Suppose that when an arbitrary orthonormal basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *is transformed to any other orthonormal basis* $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ *these numbers are changed according to the formulas*

$$L'_{ij} = \sum_{m=1}^{3}\sum_{n=1}^{3} \alpha_{im}\alpha_{jn}L_{mn}, \qquad i, j = 1, 2, 3 \tag{7.15}$$

where $\|\alpha_{ij}\|$ *is the transformation matrix from the basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *to the basis* $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$. *Then the quantity* L *is called an* ***orthogonal affine tensor of second rank (order)*** *and is denoted by the symbol* (L_{ij}), *i.e.* $\mathrm{L} \equiv (L_{ij})$.

The numbers L_{ij}, $i, j = 1, 2, 3$, are referred to as the components of the tensor L in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, the numbers L'_{ij}, $i, j = 1, 2, 3$, as its components in the basis $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ etc.

In §§ 2-9 we shall dwell in more detail on the properties of the orthogonal affine tensors of second rank and on examples of such tensors.

Now we give the definition of an orthogonal affine tensor of an arbitrary rank (order) $p \geqslant 1$.

Definition 3. Let us be given a quantity L *which is specified by a collection of* 3^p *numbers* $L_{i_1 i_2 \ldots i_p}$, $i_s = 1, 2, 3$, $s = 1, 2, \ldots, p$ *in every orthonormal basis* $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. *If these numbers are transformed according to the formulas*

$$L'_{i_1 i_2 \ldots i_p} = \sum_{j_1 j_2 \ldots j_p = 1}^{3} \alpha_{i_1 j_1} \alpha_{i_2 j_2} \ldots \alpha_{i_p j_p} L_{j_1 j_2 \ldots j_p} \qquad (7.16)$$

when an arbitrary basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ *is transformed to any other orthonormal basis* $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$ *by means of a transformation matrix* $\|\alpha_{ij}\|$, *the quantity* L *is called an* **orthogonal affine tensor of rank** p *and is denoted by the symbol* $(L_{i_1 i_2 \ldots i_p})$, *that is* $\mathrm{L} = (L_{i_1, i_2 \ldots i_p})$.

The numbers $L_{i_1, i_2 \ldots i_p}$ are called the components of the tensor L relative to the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, the numbers $L'_{i_1, i_2 \ldots i_p}$ are its components in the basis $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$ etc.

Note 1. The definition of an orthogonal affine tensor of order p ($p \geqslant 1$) can also be given in the following equivalent form.

We say that we are given an orthogonal affine tensor of rank $p \geqslant 1$ (denoted as $L_{i_1 i_2 \ldots i_p}$) if to every orthonormal basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ there correspond 3^p numbers $L_{i_1 i_2 \ldots i_p}$, $i_s = 1, 2, 3$, $s = 1, 2, \ldots, p$, which change in accord with the formulas

$$L'_{i_1 i_2 \ldots i_p} = \sum_{j_1 j_2 \ldots j_p = 1}^{3} \alpha_{i_1 j_1} \alpha_{i_2 j_2} \ldots \alpha_{i_p j_p} L_{i_1 i_2 \ldots i_p} \qquad (7.17)$$

when an arbitrary orthonormal basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is transformed to another orthonormal basis $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$ by means of a transformation matrix $\|\alpha_{ij}\|$.

We shall sometimes apply the latter form of definition to the case $p = 1$ and also to the case $p = 2$.

Note 2. Definitions 1, 2 and 3 have been given for a three-dimensional space. But they can be rephrased in a completely analogous manner for an N-dimensional space whose orthonormal bases $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N$; $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_N$ etc. contain N mutually orthogonal unit vectors, the transformation from a basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N$ to another basis $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_N$ being performed by the formulas

$$\mathbf{e}'_i = \sum_{j=1}^{N} \alpha_{ij} \mathbf{e}_j, \qquad i = 1, 2, \ldots, N$$

where the transformation matrix $\|\alpha_{ij}\|$ is of order N.

§ 2. CONNECTION BETWEEN TENSORS OF SECOND RANK AND LINEAR OPERATORS

1. Linear Operator as a Tensor of Second Rank. To begin with, we remind the reader that a **linear operator or a linear vector function**

is a function

$$y = L(x)$$

which associates, with every vector $\mathbf{x}$, a vector $\mathbf{y}$ in such a way that the relation

$$L(C_1\mathbf{x}_1 + C_2\mathbf{x}_2) = C_1 L(\mathbf{x}_1) + C_2 L(\mathbf{x}_2) \qquad (7.18)$$

are fulfilled for any $\mathbf{x}_1$ and $\mathbf{x}_2$ and any constants C_1 and C_2.

The components (coordinates) of a linear operator L in a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are the coefficients L_{ij} appearing in the resolutions of the images $L(\mathbf{e}_1)$, $L(\mathbf{e}_2)$ and $L(\mathbf{e}_3)$ of the base vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ relative to this basis:

$$\begin{aligned} L(\mathbf{e}_1) &= L_{11}\mathbf{e}_1 + L_{21}\mathbf{e}_2 + L_{31}\mathbf{e}_3 \\ L(\mathbf{e}_2) &= L_{12}\mathbf{e}_1 + L_{22}\mathbf{e}_2 + L_{32}\mathbf{e}_3 \\ L(\mathbf{e}_3) &= L_{13}\mathbf{e}_1 + L_{23}\mathbf{e}_2 + L_{33}\mathbf{e}_3 \end{aligned} \qquad (7.19)$$

These formulas can be put down, in the abbreviated notation, as

$$L(\mathbf{e}_j) = \sum_{k=1}^{3} L_{kj}\mathbf{e}_k, \qquad j = 1, 2, 3 \qquad (7.20)$$

Let us scalarly multiply both sides of equality (7.20) by the vector $\mathbf{e}_i$ ($i = 1, 2, 3$). Then, based on relations (7.3), we obtain

$$L_{ij} = (\mathbf{e}_i, L(\mathbf{e}_j)), \qquad i, j = 1, 2, 3 \qquad (7.21)$$

Similarly, for the components of the operator L in another basis $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ we have

$$L'_{ij} = (\mathbf{e}'_i, L(\mathbf{e}'_j)), \qquad i, j = 1, 2, 3 \qquad (7.22)$$

Substituting the expressions

$$\mathbf{e}'_i = \sum_{m=1}^{3} \alpha_{im}\mathbf{e}_m, \qquad \mathbf{e}'_j = \sum_{n=1}^{3} \alpha_{jn}\mathbf{e}_n \qquad (7.23)$$

into formulas (7.22) we receive the relations

$$\begin{aligned} L'_{ij} &= (\mathbf{e}'_i, L(\mathbf{e}'_j)) = \Big(\Big(\sum_{m=1}^{3} \alpha_{im}\mathbf{e}_m\Big), \Big(\sum_{n=1}^{3} \alpha_{jn} L(\mathbf{e}_n)\Big)\Big) = \\ &= \sum_{m=1}^{3}\sum_{n=1}^{3} \alpha_{im}\alpha_{jn}(\mathbf{e}_m, L(\mathbf{e}_n)) = \sum_{m=1}^{3}\sum_{n=1}^{3} \alpha_{im}\alpha_{jn} L_{mn}, \\ &\qquad i, j = 1, 2, 3 \end{aligned} \qquad (7.24)$$

Formulas (7.24) coincide with formulas (7.15) and consequently we have proved that every linear operator L is an orthogonal affine tensor of second rank.

2. Tensor of Second Rank as a Linear Operator. An orthogonal affine tensor (L_{ij}) of rank two can be interpreted as a linear operator applied to the vectors of a Euclidean space. To show this we take an orthogonal affine tensor of second rank (L_{ij}) in a three-dimensional space and define an operator $\mathbf{y} = \mathbf{L}(\mathbf{x})$ by assuming that its application to the base vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ of an arbitrary orthonormal basis is described by formulas (7.19) and that the expression $\mathbf{L}(\mathbf{x})$ (where $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ is an arbitrary vector) is specified by the formula

$$\mathbf{L}(\mathbf{x}) = \sum_{i=1}^{3} x_i \mathbf{L}(\mathbf{e}_i) \tag{7.25}$$

Thus, we have associated, with every orthogonal tensor (L_{ij}) of second rank, a certain vector function (operator) $\mathbf{y} = \mathbf{L}(\mathbf{x})$. Now we are going to prove that the operator thus defined is in fact linear, i.e. satisfies condition (7.18). Let $\mathbf{x} = \sum_{i=1}^{3} x_i\mathbf{e}_i$ and $\mathbf{y} = \sum_{i=1}^{3} y_i\mathbf{e}_i$; then we have

$$C_1\mathbf{x} + C_2\mathbf{y} = \sum_{i=1}^{3} (C_1x_i + C_2y_i)\,\mathbf{e}_i$$

Consequently, by virtue of definition (7.25), we obtain (substituting $C_1\mathbf{x} + C_2\mathbf{y}$ for $\mathbf{x}$ into (7.25)) the relation

$$\mathbf{L}(C_1\mathbf{x} + C_2\mathbf{y}) = \sum_{i=1}^{3} (C_1x_i + C_2y_i)\,\mathbf{L}(\mathbf{e}_i) =$$

$$= C_1 \sum_{i=1}^{3} x_i\mathbf{L}(\mathbf{e}_i) + C_2 \sum_{i=1}^{3} y_i\mathbf{L}(\mathbf{e}_i) = C_1\mathbf{L}(\mathbf{x}) + C_2\mathbf{L}(\mathbf{y})$$

which coincides with (7.18). Hence, the linearity of the operator has been proved.

It can be easily shown that the linear operator $\mathbf{L}$ defined above by means of the tensor (L_{ij}) does not depend on the choice of the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. In other words, if instead of the components L_{ij} of the tensor in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ we take its components L'_{ij} in another basis $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ and define a linear operator $\mathbf{L}'$ by means of the relations

$$\mathbf{L}'(\mathbf{e}'_i) = \sum_{k=1}^{3} L'_{ki}\mathbf{e}'_k \quad (i = 1, 2, 3), \qquad \mathbf{L}'(\mathbf{x}) = \sum_{i=1}^{3} x'_i\mathbf{L}'(\mathbf{e}'_i) \tag{7.25'}$$

where $\mathbf{x} = \sum_{i=1}^{3} x'_i\mathbf{e}'_i$ we shall have the identity

$$\mathbf{L}'(\mathbf{x}) \equiv \mathbf{L}(\mathbf{x}) \tag{7.26}$$

for each vector $\mathbf{x}$.

Indeed, taking advantage of the relations

$$L'_{lj} = \sum_{k=1}^{3} \sum_{i=1}^{3} \alpha_{lk}\alpha_{ji} L_{ki}, \quad x_i = \sum_{j=1}^{3} \alpha_{ji} x'_j, \quad \mathbf{e}_k = \sum_{l=1}^{3} \alpha_{lk} \mathbf{e}'_l$$

and formulas (7.19), (7.25) and (7.25′) we obtain

$$\begin{aligned} \mathrm{L}(\mathbf{x}) &= \sum_{i=1}^{3} x_i \mathrm{L}(\mathbf{e}_i) = \sum_{i=1}^{3} x_i \sum_{k=1}^{3} L_{ki}\mathbf{e}_k = \\ &= \sum_{j=1}^{3} x'_j \sum_{l=1}^{3} \Big(\sum_{k=1}^{3} \sum_{i=1}^{3} \alpha_{lk}\alpha_{ji} L_{ki} \Big) \mathbf{e}'_l = \sum_{j=1}^{3} x'_j \sum_{l=1}^{3} L'_{lj}\mathbf{e}'_l = \\ &= \sum_{j=1}^{3} x'_j \mathrm{L}'(\mathbf{e}'_j) = \mathrm{L}'(\mathbf{x}) \end{aligned} \tag{7.27}$$

which is what we set out to prove.

We have thus shown that the operators L′ and L coincide and consequently relations (7.19) and (7.25) establish a one-to-one correspondence between orthogonal affine tensors (L_{ij}) of rank two and the linear operators L associated with them. The linear operator L can be identified with the corresponding tensor (L_{ij}), and hence we can consider an orthogonal affine tensor of second rank as being a linear operator. This interpretation of an orthogonal affine tensor of second rank is widely used in physics. Namely, in this manner we interpret the *conductivity tensor* mentioned at the beginning of the present chapter. The *inertia tensor* introduced in mechanics and the *stress tensor* considered in the theory of elasticity (the latter tensor will be treated in § 5*) are also understood in this way. But there is another interpretation of a tensor of second rank which proves to be useful in various applications. This new approach to the notion of tensor is discussed in § 3.

§ 3. CONNECTION BETWEEN TENSORS AND INVARIANT MULTILINEAR FORMS

1. Tensors of Rank One and Invariant Linear Forms. Let us be given, in every coordinate system, a triple of numbers a_1, a_2, a_3. Suppose that when one coordinate system is transformed

* The reader will see in § 5 that the interpretation of the stress tensor is connected with the notion of the *adjoint operator*. If, instead of relations (7.20) and (7.25) specifying the linear operator L corresponding to a tensor (L_{ij}), we take the formulas $\mathrm{L}^*(\mathbf{e}_j) = \sum_{k=1}^{3} L_{jk}\mathbf{e}_k$, $j = 1, 2, 3$, and $\mathrm{L}^*(\mathbf{x}) = \sum_{i=1}^{3} x_i \mathrm{L}^*(\mathbf{e}_i)$ $\left(\mathbf{x} = \sum_{i=1}^{3} x_i\mathbf{e}_i\right)$ we arrive at the linear operator L* which is the adjoint of L.

to another these numbers change in such a way that the linear form $a_1x_1 + a_2x_2 + a_3x_3$ (where x_1, x_2 and x_3 are the coordinates of an arbitrary vector **x**) remains invariant. Then it turns out that the quantities a_i ($i = 1, 2, 3$) form a tensor of rank one. Indeed, let an arbitrary vector **x** be resolved as $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ relative to a basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ in which the coefficients of a given linear form are a_1, a_2, a_3, and let the same vector **x** be expressed in the form $\mathbf{x} = x_1'\mathbf{e}_1' + x_2'\mathbf{e}_2' + x_3'\mathbf{e}_3'$ in another basis $\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'$ for which the coefficients of the linear form are a_1', a_2' and a_3'. The linear form being regarded as invariant, we thus have the equality

$$a_1'x_1' + a_2'x_2' + a_3'x_3' = a_1x_1 + a_2x_2 + a_3x_3 \tag{7.28}$$

for every vector **x**. Let us substitute the expression of x_k in terms of x_i' (i.e. $x_k = \sum_{i=1}^{3} \alpha_{ik}x_i'$) into the right-hand side of equality (7.28). This results in

$$\sum_{i=1}^{3} a_i' x_i' = \sum_{k=1}^{3} a_k \sum_{i=1}^{3} \alpha_{ik} x_i' = \sum_{i=1}^{3} \Big(\sum_{k=1}^{3} \alpha_{ik} a_k \Big) x_i'$$

The quantities x_1', x_2' and x_3' being quite arbitrary here, we can write the relation

$$a_i' = \sum_{k=1}^{3} \alpha_{ik} a_k \tag{7.29}$$

which is what we set out to prove.

2. **Tensors of Rank Two and Invariant Bilinear Forms.** We can similarly prove that the coefficients of an invariant bilinear form

$$\sum_{i,\, j=1}^{3} a_{ij}x_iy_j \tag{7.30}$$

(where x_i and y_i, $i = 1, 2, 3$, are, respectively, the coordinates of variable vectors **x** and **y**) constitute a tensor of second rank. In fact, let the bilinear form be expressed as (7.30) in a basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and as

$$\sum_{i,\, j=1}^{3} a_{ij}' x_i' y_j' \tag{7.31}$$

in another basis $\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'$. We suppose the form to be invariant, and consequently

$$\sum_{i,\, j=1}^{3} a_{ij}' x_i' y_j' = \sum_{m,\, n=1}^{3} a_{mn} x_m y_n \tag{7.32}$$

for any two vectors **x** and **y**. Substituting the expressions

$$x_m = \sum_{i=1}^{3} \alpha_{im} x_i', \quad y_n = \sum_{j=1}^{3} \alpha_{jn} y_j' \quad (m, n = 1, 2, 3) \tag{7.33}$$

of the old coordinates of the vectors x and y (i.e. their coordinates relative to the basis e_1, e_2, e_3) in terms of the new coordinates (in the basis e'_1, e'_2, e'_3) into the right-hand side of equality (7.32) we deduce the relation

$$\sum_{i, j=1}^{3} a'_{ij} x'_i y'_j = \sum_{m, n=1}^{3} a_{mn} \left(\sum_{i=1}^{3} \alpha_{im} x'_i\right) \left(\sum_{j=1}^{3} \alpha_{jn} y'_j\right) =$$
$$= \sum_{i, j=1}^{3} \left(\sum_{m, n=1}^{3} \alpha_{im}\alpha_{jn}a_{mn}\right) x'_i y'_j \tag{7.34}$$

The arbitrariness of x'_i and y'_i ($i = 1, 2, 3$) suggests that

$$a'_{ij} = \sum_{m, n=1}^{3} \alpha_{im}\alpha_{jn}a_{mn} \tag{7.35}$$

which is what we set out to prove.

Note 1. Equality (7.35) can be proved by considering only vector of unit length. Indeed, putting

$$x'_i = \begin{cases} 1 \text{ for } i = i_0 \\ 0 \text{ for } i \neq i_0 \end{cases} \quad i = 1, 2, 3; \quad y'_j = \begin{cases} 1 \text{ for } j = j_0 \\ 0 \text{ for } j \neq j_0 \end{cases} \quad j = 1, 2, 3 \tag{7.36}$$

we derive from equality (7.34) the relation

$$a'_{i_0 j_0} = \sum_{m, n=1}^{3} \alpha_{i_0 m}\alpha_{j_0 n}a_{mn} \qquad (i_0, j_0 = 1, 2, 3)$$

Furthermore, by equality (7.36), the vectors

$$\mathbf{x} = \sum_{i=1}^{3} x'_i \mathbf{e}'_i \quad \text{and} \quad \mathbf{y} = \sum_{j=1}^{3} y'_j \mathbf{e}'_j$$

are of unit length because the basis e'_1, e'_2, e'_3 is orthonormal (w have agreed to restrict ourselves to such bases). Consequently, a bilinear form is invariant on unit sphere, that is on conditio that its values are considered only for the vectors of unit lengtl its coefficients a_{ij} ($i, j = 1, 2, 3$) constitute an orthogonal affi tensor of second rank.

A bilinear form is said to be symmetric if its coefficient matri is symmetric, that is if $a_{ij} = a_{ji}$, $i, j = 1, 2, 3$. (By virtue relation (7.35), we can assert that if $a_{ij} = a_{ji}$ at least in one orth normal basis the matrix $\| a_{ij} \|$ remains symmetric for any oth orthonormal basis.) Putting $\mathbf{y} = \mathbf{x}$ in a symmetric bilinear for we obtain the so-called quadratic form

$$\sum_{i, j=1}^{3} a_{ij} x_i y_j \tag{7.3}$$

A symmetric bilinear form is uniquely specified by the quadratic form generated by it when we put $\mathbf{x} = \mathbf{y}$. Actually, if we substitute the coordinates of the vector $\mathbf{x} + \mathbf{y}$ for those of an arbitrary vector $\mathbf{x}$ (where $\mathbf{y}$ is also an arbitrary vector) into formula (7.37) we obtain

$$\sum_{j=1}^{3} a_{ij}(x_i + y_i)(x_j + y_j) = \sum_{i,\, j=1}^{3} a_{ij}x_ix_j + \sum_{i,\, j=1}^{3} a_{ij}y_iy_j +$$
$$+ \sum_{i,\, j=1}^{3} a_{ij}x_iy_j + \sum_{i,\, j=1}^{3} a_{ij}y_ix_j =$$
$$= \sum_{i,\, j=1}^{3} a_{ij}x_iy_j + \sum_{i,\, j=1}^{3} a_{ij}y_iy_j + 2\sum_{i,\, j=1}^{3} a_{ij}x_iy_j \tag{7.38}$$

since $a_{ij} = a_{ji}$. Consequently, we receive the equality

$$\sum_{i,\, j=1}^{3} a_{ij}x_iy_j =$$
$$= \frac{1}{2}\left\{\sum_{i,\, j=1}^{3} a_{ij}(x_i + y_i)(x_j + y_j) - \sum_{i,\, j=1}^{3} a_{ij}x_ix_j - \sum_{i,\, j=1}^{3} a_{ij}y_iy_j\right\} \tag{7.39}$$

which is what we set out to prove.

It follows that the coefficients of an invariant quadratic form constitute an orthogonal affine tensor of second rank. Indeed, they coincide with the coefficients of the corresponding invariant symmetric bilinear form for which we have already proved that the collection of its coefficients is an orthogonal affine tensor of second rank.

Note 2. On the basis of Note 1 we conclude that the coefficients a_{ij} $(a_{ij} = a_{ji};\ i, j = 1, 2, 3)$ of a quadratic form $\sum_{i,\, j=1}^{3} a_{ij}x_ix_j$ which is defined and invariant on unit sphere constitute an orthogonal affine tensor of second rank.

3. Tensors of Arbitrary Rank p and Invariant Multilinear Forms. Let vectors $\xi_1, \xi_2, \ldots, \xi_p$ be resolved with respect to a basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$:

$$\xi_j = \xi_{j1}\mathbf{e}_1 + \xi_{j2}\mathbf{e}_2 + \xi_{j3}\mathbf{e}_3, \quad j = 1, 2, \ldots, p$$

Suppose that, for every basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, there is a system of coefficients $a_{i_1i_2\ldots i_p}$ (where $i_s = 1, 2, 3,\ \ s = 1, 2, \ldots, p$). Then the function

$$\sum_{i_1, i_2, \ldots, i_p=1}^{3} a_{i_1i_2\ldots i_p}\xi_{1i_1}\xi_{2i_2}\ldots\xi_{pi_p}$$

is said to be a multilinear form. As in the case of an invariant bilinear form, it can be easily proved that the collection of the coefficients of an invariant multilinear form of an arbitrary order $p \geqslant 1$ is an orthogonal affine tensor of rank p.

§ 4. STRAIN TENSOR

Consider a deformable physical body whose arbitrary point is specified by its position vector (radius vector) $\mathbf{r} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ in a coordinate system $Ox_1x_2x_3$. If the radius vector of a point M is equal to $\mathbf{r}$ we have $\overline{OM} = \mathbf{r}$. In this case we shall write $M(\mathbf{r})$.

Suppose the body is subjected to a deformation in which a point $M(\mathbf{r})$ is displaced by a vector $\mathbf{u}$, that is passes to the new position $M'(\mathbf{r} + \mathbf{u})$ (see Fig. 7.1). The deformation is specified by the field

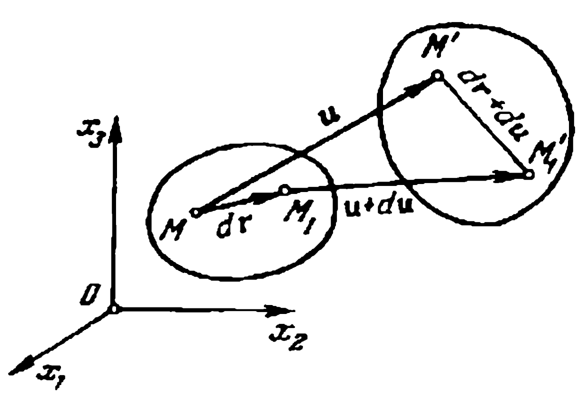


Fig. 7.1

of displacements $\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3$. Let us consider a point $M_1(\mathbf{r} + d\mathbf{r})$ lying close to the point $M(\mathbf{r})$. After the body has been deformed this point occupies the position $M_1'(\mathbf{r} + d\mathbf{r} + \mathbf{u} + d\mathbf{u})$. The deformation of the body in the vicinity of the point $M(\mathbf{r})$ can be characterized by the variations of the lengths of all the line segments $\overline{MM_i}$ $(i = 1, 2, 3, \ldots)$ starting from this point, their end points $M_1, M_2, \ldots$ lying in a sufficiently small neighbourhood of the point $M(\mathbf{r})$.

Let us investigate the variation of the length of the line segment $\overline{MM_1}$ due to the deformation of the body. The length of the segment $\overline{MM_1}$ in its original position is equal to $|d\mathbf{r}|$. The segment will occupy, after the deformation, the position of the line segment $\overline{M'M_1'}$ whose length is equal to $|d\mathbf{r} + d\mathbf{u}|$. As a measure of change of the length of the segment $\overline{MM_1}$, we shall take the quantity

$$\frac{1}{2}\{\overline{M'M_1'}^2 - \overline{MM_1}^2\} = \frac{1}{2}\{(d\mathbf{r} + d\mathbf{u})^2 - d\mathbf{r}^2\} = \frac{1}{2}\{2d\mathbf{u}\,d\mathbf{r} + d\mathbf{u}^2\} =$$

$$= \gamma_{x_1x_1}\,dx_1^2 + \gamma_{x_2x_2}\,dx_2^2 + \gamma_{x_3x_3}\,dx_3^2 + 2\gamma_{x_1x_2}\,dx_1\,dx_2 + 2\gamma_{x_1x_3}\,dx_1\,dx_3 + 2\gamma_{x_2x_3}\,dx_2\,dx_3$$

where $\gamma_{x_i x_j}$ $(i, j = 1, 2, 3)$ are some coefficients and dx_i $(i = 1, 2, 3)$ are the components of $d\mathbf{r}$. This expression is a quadratic form in the variables dx_1, dx_2 and dx_3. The construction of this form implies that it is invariant. Consequently, its coefficients constitute a tensor of rank two which is characterized by the matrix

$$\begin{Vmatrix} \gamma_{x_1x_1} & \gamma_{x_1x_2} & \gamma_{x_1x_3} \\ \gamma_{x_2x_1} & \gamma_{x_2x_2} & \gamma_{x_2x_3} \\ \gamma_{x_3x_1} & \gamma_{x_3x_2} & \gamma_{x_3x_3} \end{Vmatrix} \qquad (7.40)$$

The tensor thus formed is called the **strain tensor**.

Suppose that the deformation is so small that the squares and the products of the derivatives of u_1, u_2 and u_3 with respect to x_1, x_2 and x_3 are negligibly small compared with the first-order terms. Then the matrix of the strain tensor can be written in the form

$$\begin{Vmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{1}{2}\left(\dfrac{\partial u_1}{\partial x_2} + \dfrac{\partial u_2}{\partial x_1}\right) & \dfrac{1}{2}\left(\dfrac{\partial u_1}{\partial x_3} + \dfrac{\partial u_3}{\partial x_1}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial u_2}{\partial x_1} + \dfrac{\partial u_1}{\partial x_2}\right) & \dfrac{\partial u_2}{\partial x_2} & \dfrac{1}{2}\left(\dfrac{\partial u_2}{\partial x_3} + \dfrac{\partial u_3}{\partial x_2}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial u_3}{\partial x_1} + \dfrac{\partial u_1}{\partial x_3}\right) & \dfrac{1}{2}\left(\dfrac{\partial u_3}{\partial x_2} + \dfrac{\partial u_2}{\partial x_3}\right) & \dfrac{\partial u_3}{\partial x_3} \end{Vmatrix} \qquad (7.41)$$

§ 5. STRESS TENSOR

1. Definition of Stress Tensor. Suppose we have an elastic body which has been deformed. Let us mentally draw an elementary plane surface of area σ through a point M of the body and erect at M a unit normal vector $\mathbf{n}$ to one of the two sides of the surface (see Fig. 7.2). If we divide the resultant elastic force $\mathbf{F}_{n\sigma}$ (applied to the chosen side of the surface element) by the area σ we obtain the so-called *average (mean) stress* $(\mathbf{p}_n)_{av} = \frac{\mathbf{F}_{n\sigma}}{\sigma}$ on the elementary area σ with normal $\mathbf{n}$ drawn through the point M. Passing to the limit as σ is contracted toward the point M we arrive at the (*total*) *stress* $\mathbf{p}_n$ at the point M on an elementary area with normal $\mathbf{n}$:

$$\mathbf{p}_n = \lim_{\sigma \to M} \frac{\mathbf{F}_{n\sigma}}{\sigma}. \qquad (7.42)$$

Changing the direction of the normal $\mathbf{n}$, that is turning the area σ about the point M it is drawn through, we obtain different values of the vector $\mathbf{p}_n$ at the same point M. Thus, a state of stress of an elastic body at a given point M cannot be completely characterized by a single vector. But it turns out that to obtain an exhaustive description of such a state it is sufficient to determine the stresses on three mutually perpendicular plane sections passing through the point M because this makes it possible to find the stress at the point M on an area (passing through M) of arbitrary orientation.

Let us establish the result stated above. Denote by $\mathbf{p}_{x_1}$, $\mathbf{p}_{x_2}$ and $\mathbf{p}_{x_3}$ the stresses at the point M on three elementary areas whose normals go in the positive directions of the coordinate axes Ox_1, Ox_2 and Ox_3. In other words, $\mathbf{p}_{x_i}$ is the stress on the area with unit normal $\mathbf{e}_i$, $i = 1, 2, 3$, where $\mathbf{e}_i$ is the base vector of the axis Ox_i (Fig. 7.3). Consider a tetrahedron with one vertex at the point M and the edges MA, MB and MC parallel to the axes Ox_1, Ox_2 and Ox_3. The outer normal $\mathbf{n}_2$ to the face MAC of the tetrahedron and the vector $\mathbf{e}_2$ are in the opposite directions. Hence, the

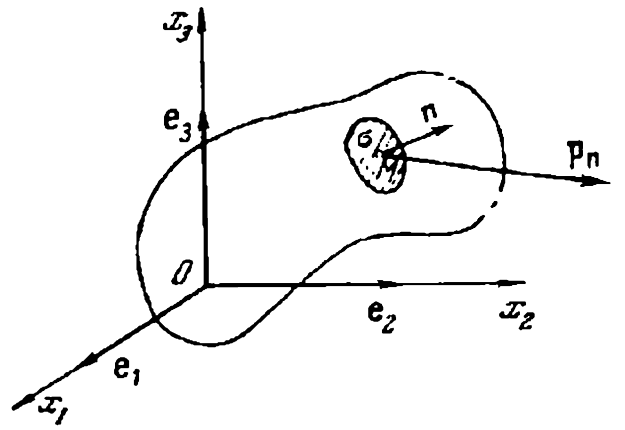


Fig. 7.2

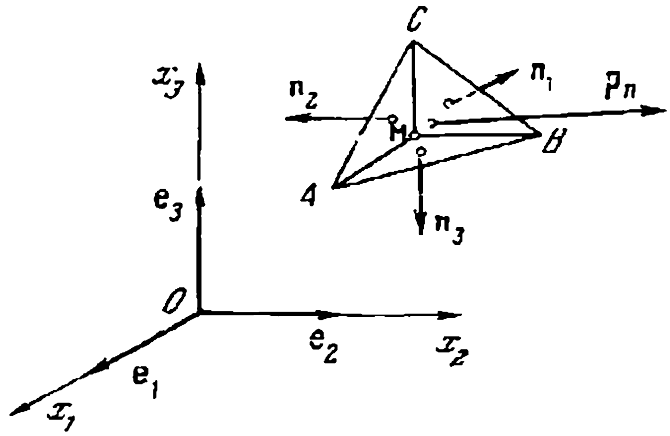


Fig. 7.3

stress on the area MAC is equal to $-\mathbf{p}_{x_2}$. Similarly, the stress on the face BMC whose outer normal is $\mathbf{n}_1 = -\mathbf{e}_1$ is equal to $-\mathbf{p}_{x_1}$, and the stress on the area MAB corresponding to the outer normal $\mathbf{n}_3 = -\mathbf{e}_3$ is equal to $\mathbf{p}_{x_3}$. Let us designate by $\mathbf{p}_n$ the stress on the area ABC with outer normal $\mathbf{n}$ and form the equation expressing Newton's second law for the tetrahedron $MABC$:

$$\rho \frac{1}{3} \sigma h \frac{d\mathbf{v}}{dt} = \sigma \mathbf{p}_n - \sigma \cos(\mathbf{n}, x_1)\, \mathbf{p}_{x_1} - \\ - \sigma \cos(\mathbf{n}, x_2)\, \mathbf{p}_{x_2} - \sigma \cos(\mathbf{n}, x_3)\, \mathbf{p}_{x_3} + \frac{1}{3} \sigma h \rho \mathbf{f} \qquad (7.43)$$

Here $\frac{dv}{dt}$ is the acceleration, σ is the area of the face ABC, h is the altitude of the tetrahedron (if the face ABC is taken as its base), ρ is the volume mass density, $\mathbf{f}$ the volume force per unit mass (in particular, it may be the gravity force), the quantity $\frac{1}{3}\sigma h$ is the volume of the tetrahedron $MABC$ and the quantities $\sigma \cos(\mathbf{n}, x_1)$, $\sigma \cos(\mathbf{n}, x_2)$ and $\sigma \cos(\mathbf{n}, x_3)$ are, respectively, the areas of the faces MBC, MAC and MAB. If we suppose that the acceleration $\frac{dv}{dt}$ and the volume force $\mathbf{f}$ remain bounded when σ is made to tend to zero, then dividing equality (7.43) by σ we obtain, in the limit, the relation

$$0 = \mathbf{p}_n - \mathbf{p}_{x_1} \cos(\mathbf{n}, x_1) - \mathbf{p}_{x_2} \cos(\mathbf{n}, x_2) - \mathbf{p}_{x_3} \cos(\mathbf{n}, x_3)$$

Consequently,

$$\mathbf{p}_n = \mathbf{p}_{x_1}\cos(\mathbf{n}, x_1) + \mathbf{p}_{x_2}\cos(\mathbf{n}, x_2) + \mathbf{p}_{x_3}\cos(\mathbf{n}, x_3) \qquad (7.44)$$

Formula (7.44) expresses the stress $\mathbf{p}_n$ on an area with an arbitrary normal $\mathbf{n}$ in terms of the stresses on the areas whose normals go along the coordinate axes.

Let us resolve the vectors $\mathbf{p}_{x_1}$, $\mathbf{p}_{x_2}$ and $\mathbf{p}_{x_3}$ along the base vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$\begin{aligned} \mathbf{p}_{x_1} &= p_{11}\mathbf{e}_1 + p_{12}\mathbf{e}_2 + p_{13}\mathbf{e}_3 \\ \mathbf{p}_{x_2} &= p_{21}\mathbf{e}_1 + p_{22}\mathbf{e}_2 + p_{23}\mathbf{e}_3 \\ \mathbf{p}_{x_3} &= p_{31}\mathbf{e}_1 + p_{32}\mathbf{e}_2 + p_{33}\mathbf{e}_3 \end{aligned} \qquad (7.45)$$

If, for a given point M, the matrix

$$\| p_{ij} \| = \begin{Vmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{Vmatrix} \qquad (7.46)$$

is known, we can determine the stress on any area σ passing through the point M because the position of the area is specified by the direction of its normal $\mathbf{n}$ and formulas (7.44) and (7.45) make it possible to find $\mathbf{p}_n$ if the vector $\mathbf{n}$ is given. Thus, a state of stress of an elastic body at a given point is completely characterized by matrix (7.46).

Consider now the projection of the vector $\mathbf{p}_n$ on the normal $\mathbf{n}$. The physical significance of the quantity thus obtained suggests that it is independent of the choice of the coordinate system. To find the projection we scalarly multiply both sides of equality (7.44) by $\mathbf{n}$ and apply formulas (7.45). This yields the expression

$$(\mathbf{p}_n, \mathbf{n}) = \sum_{i, j=1}^{n} p_{ij}\cos(\mathbf{n}, x_i)\cos(\mathbf{n}, x_j) \qquad (7.47)$$

and hence the sought-for quantity is given by a quadratic form defined on unit sphere.

Hence, quadratic form (7.47) is invariant on unit sphere and therefore its coefficients p_{ij} $(i, j = 1, 2, 3)$ constitute an orthogonal affine tensor of second rank $\Pi = (p_{ij})$ (see Note 2 in § 3). This is the so-called stress tensor.

2. **Stress Tensor as a Linear Operator.** It is convenient to interpret the stress tensor as a linear operator transforming the unit normal $\mathbf{n}$ to an area into the vector $\mathbf{p}_n$ (the total stress on the area).

Let us take the resolution of the vector $\mathbf{p}_n$, with respect to the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, which is of the form

$$\mathbf{p}_n = p_{n_1}\mathbf{e}_1 + p_{n_2}\mathbf{e}_2 + p_{n_3}\mathbf{e}_3$$

and substitute it into the left-hand side of equality (7.44) and simultaneously substitute the expressions of the vectors $\mathbf{p}_{x_1}$, $\mathbf{p}_{x_2}$ and $\mathbf{p}_{x_3}$ in terms of the same basis (see relations (7.45)) into the right-hand side of the equality. The resolution of $\mathbf{p}_n$ with respect to the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ being unique, we thus obtain the following system of three scalar equations:

$$\left.\begin{aligned} p_{n1} &= p_{11}\cos(\mathbf{n}, x_1) + p_{21}\cos(\mathbf{n}, x_2) + p_{31}\cos(\mathbf{n}, x_3)\\ p_{n2} &= p_{12}\cos(\mathbf{n}, x_1) + p_{22}\cos(\mathbf{n}, x_2) + p_{32}\cos(\mathbf{n}, x_3)\\ p_{n3} &= p_{13}\cos(\mathbf{n}, x_1) + p_{23}\cos(\mathbf{n}, x_2) + p_{33}\cos(\mathbf{n}, x_3)\end{aligned}\right\} \qquad (7.48)$$

These equations express the above mentioned linear operator in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. The unit normal vector $\mathbf{n}$ to an area σ is expressed as

$$\mathbf{n} = \mathbf{e}_1\cos(\mathbf{n}, x_1) + \mathbf{e}_2\cos(\mathbf{n}, x_2) + \mathbf{e}_3\cos(\mathbf{n}, x_3)$$

Presenting the vector $\mathbf{n}$ as a row matrix of the form

$$\mathbf{n} = \|\cos(\mathbf{n}, x_1), \cos(\mathbf{n}, x_2), \cos(\mathbf{n}, x_3)\|$$

we can write (see Appendix to Chapter 7)

$$\mathbf{p}_n = \mathbf{n}\,\|p_{ij}\| \qquad (7.49)$$

where $\mathbf{p}_n$ is interpreted as a row matrix $\|p_{n1}, p_{n2}, p_{n3}\|$ and $\|p_{ij}\|$ is the matrix corresponding to the stress tensor $\Pi = (p_{ij})$ at the given point. When speaking about the multiplication of a matrix corresponding to a tensor by a vector we simply say that the tensor is multiplied by the vector.* Thus, to obtain the total stress at a point M on an area σ with unit normal $\mathbf{n}$ we must take the stress tensor $\Pi = (p_{ij})$ at the point and multiply it on the left by the vector $\mathbf{n}$:

$$\mathbf{p}_n = \mathbf{n}\,(p_{ij}) = \mathbf{n}\Pi \qquad (7.49')$$

Of course, we can also represent the unit normal $\mathbf{n}$ as a column matrix

$$\mathbf{n} = \begin{Vmatrix}\cos(\mathbf{n}, x_1)\\ \cos(\mathbf{n}, x_2)\\ \cos(\mathbf{n}, x_3)\end{Vmatrix}$$

Then $\mathbf{p}_n$ (understood as a column matrix $\begin{Vmatrix}p_{n1}\\ p_{n2}\\ p_{n3}\end{Vmatrix}$) can be obtained by multiplying on the right the transpose of the matrix $\|p_{ij}\|$

* More precisely, we speak here about the multiplication of a tensor (interpreted as a linear operator) by a vector in the sense that the corresponding linear operator is applied to the vector. See also the end of this section.

by the above vector $\mathbf{n}$. This corresponds to the application of the adjoint operator (see footnote on page 291) associated with the tensor Π to the vector (column matrix) $\mathbf{n}$.

§ 6. ALGEBRAIC OPERATIONS ON TENSORS

1. Addition, Subtraction and Multiplication of Tensors. The operations of addition and subtraction can be performed on tensors of the same rank. For instance, the *sum* of two tensors of second rank a_{ij} and b_{ij} is the tensor c_{ij} whose components are

$$c_{ij} = a_{ij} + b_{ij}; \quad i, j = 1, 2, 3$$

and their *difference* is the tensor $d_{ij} = a_{ij} - b_{ij}$ $(i, j = 1, 2, 3)$. It can be easily shown that the quantities $c_{ij} = a_{ij} + b_{ij}$ and $d_{ij} = a_{ij} - b_{ij}$ are transformed according to the rule of transformation of tensors when the coordinate system is changed. The addition and subtraction of two tensors of an arbitrary rank are defined similarly.

The *product* of tensors (also spoken of as the *outer product*) can be defined for tensors of any rank. For example, the product of a tensor of rank two a_{ij} by a tensor of rank three b_{mnp} is a tensor of rank five whose components c_{ijmnp} are defined by the relation

$$c_{ijmnp} = a_{ij}b_{mnp}; \quad i, j, m, n, p = 1, 2, 3$$

We can easily prove that the above quantities c_{ijmnp} are transformed in accord with the rule of transformation of tensors when we pass from one coordinate system to another. The (outer) product of two tensors of arbitrary ranks is defined similarly.

The product of a tensor by a number can be considered a special case of the product of two tensors and is defined as follows: the product of a tensor a_{ijk} by a number C is the tensor with the components $b_{ijk} = Ca_{ijk}$. The fact that the quantities b_{ijk} constitute a tensor can be easily verified.

2. Multiplying Tensor by Vector. When studying the stress tensor (see relation (7.49) in § 5), we dealt with a special case of multiplying a tensor of second rank (interpreted as a linear operator) by a vector. Here we shall discuss this operation in the general form.

A tensor (L_{ij}) can be multiplied by a vector $\mathbf{x}$ *on the left* or *on the right*, i.e. we distinguish between the products $\mathbf{x}\,(L_{ij})$ and $(L_{ij})\,\mathbf{x}$. But in both cases the result of the operation is a vector which is defined as follows. Let the matrix corresponding to the tensor (L_{ij}) in the coordinate system specified by a basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be equal to

$$\| L_{ij} \| = \begin{Vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{Vmatrix} \tag{7.50}$$

and let the resolution of $\mathbf{x}$ relative to the basis be of the form $\mathbf{x} = \sum_{i=1}^{3} x_i \mathbf{e}_i$. Then the vector

$$\mathbf{y}^* = \mathbf{x}\,(L_{ij}) \tag{7.51}$$

resulting from the multiplication on the left of the tensor (L_{ij}) by the vector $\mathbf{x}$ is regarded as a row matrix $\| y_1^*, y_2^*, y_3^* \|$ whose components y_1^*, y_2^* and y_3^* are determined in this basis by the relation (see Appendix to Chapter 7)

$$\| y_1^*, y_2^*, y_3^* \| = \| x_1, x_2, x_3 \| \begin{Vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{Vmatrix} \tag{7.52}$$

The vector

$$\mathbf{y} = (L_{ij})\,\mathbf{x} \tag{7.53}$$

resulting from the multiplication on the right of (L_{ij}) by $\mathbf{x}$ is given by the relation

$$\begin{Vmatrix} y_1 \\ y_2 \\ y_3 \end{Vmatrix} = \begin{Vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} \tag{7.54}$$

specifying $\mathbf{y} = \sum_{i=1}^{3} y_i \mathbf{e}_i$ in the same basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Relations (7.52) and (7.54) can be written in the abridged notation

$$\mathbf{y}^* = \mathbf{x}\, \| L_{ij} \| \tag{7.52'}$$

and

$$\mathbf{y} = \| L_{ij} \|\, \mathbf{x} \tag{7.54'}$$

The vectors $\mathbf{x}$ and $\mathbf{y}^*$ entering into the former relation are interpreted as row matrices and the vectors $\mathbf{x}$ and $\mathbf{y}$ in the latter relation as column matrices.*

3. **Contraction.** *Contraction* of an orthogonal affine tensor is the operation of putting one index equal to another and then summing with respect to that index. For instance, if we take a tensor of rank four c_{ijmn}, put $i = j$ and sum with respect to i we obtain the contracted tensor of second rank whose components are given by the equalities

$$a_{mn} = \sum_{i=1}^{3} c_{iimn}$$

The fact that the quantities a_{mn} constitute a tensor of second rank can be easily proved. If we take a tensor of an even rank and perform

* Relations (7.51) and (7.53) (or, which is the same, relations (7.52') and (7.54')) determine two linear operators which are the adjoints of each other (see footnote on page 291).

contraction operations on it as many times as possible we shall arrive at a scalar, i.e. an invariant.

For example, if the contraction is performed on the product $a_i b_j$ of two tensors a_i and b_j of rank one the resultant tensor will be the invariant scalar which is nothing but the scalar product of the vector **a** (with components a_1, a_2 and a_3) by the vector **b** (whose components are b_1, b_2 and b_3):

$$(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{3} a_i b_i$$

4. Interchanging Indices. Let us consider the operation of *interchanging indices* for an important special case, namely for an orthogonal affine tensor of second rank (L_{ij}). Taking an arbitrary orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in which the components of the tensor are L_{ij} $(i, j = 1, 2, 3)$ we put

$$L_{ij}^* = L_{ji}$$

and thus arrive at the quantities L_{ij}^* $(i, j = 1, 2, 3)$ specified in every orthonormal basis. It can be easily proved that L_{ij}^* form an orthogonal affine tensor of second rank. The tensor (L_{ij}^*) constituted by the quantities L_{ij}^* $(i, j = 1, 2, 3)$ is called the **conjugate tensor** of (L_{ij}). The above operation is similarly performed on a tensor of an arbitrary rank in which any two indices can be interchanged. The resultant tensor obviously has the same rank.

5. Resolution of Tensor of Second Rank into Symmetric and Antisymmetric Parts. An orthogonal affine tensor of second rank (L_{ij}) is said to be symmetric if its matrix

$$\| L_{ij} \| = \begin{Vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{Vmatrix}$$

is symmetric in each orthonormal basis, that is if in every such basis the relations $L_{ij} = L_{ji}$ $(i, j = 1, 2, 3)$ hold.

A tensor of rank two (L_{ij}) is called **antisymmetric (skew-symmetric)** if the elements of the matrix $\| L_{ij} \|$ corresponding to it satisfy the conditions

$$L_{ij} = -L_{ji}$$

in every orthonormal basis. The latter relations suggest that for an antisymmetric tensor (L_{ij}) we always have $L_{ii} = -L_{ii}$, i.e. $2L_{ii} = 0$ and $L_{ii} = 0$.

Thus, a symmetric tensor of second rank is completely specified by its six components (since $L_{12} = L_{21}$, $L_{13} = L_{31}$ and $L_{23} = L_{32}$ for such a tensor) whereas an antisymmetric tensor is characterized by its three nondiagonal elements.

The vector product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is a simple example of an antisymmetric tensor. Indeed, let the vectors $\mathbf{a}$ and $\mathbf{b}$ have the resolutions

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 \quad \text{and} \quad \mathbf{b} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + b_3\mathbf{e}_3$$

in a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Then the vector product $[\mathbf{a}, \mathbf{b}]$ can be written as

$$[\mathbf{a}, \mathbf{b}] = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} =$$

$$= (a_2b_3 - a_3b_2)\,\mathbf{e}_1 + (a_3b_1 - a_1b_3)\,\mathbf{e}_2 + (a_1b_2 - a_2b_1)\,\mathbf{e}_3 \qquad (7.55)$$

Taking advantage of the fact that the vectors $\mathbf{a}$ and $\mathbf{b}$ are tensors of rank one we can easily show that the nine quantities $L_{ij} = a_ib_j - a_jb_i$ $(i, j = 1, 2, 3)$ form a tensor of second rank. This tensor is obviously antisymmetric because we have $L_{ji} = a_jb_i - a_ib_j = -(a_ib_j - a_jb_i) = -L_{ij}$. Consequently, the tensor is completely specified by its three components $(a_2b_3 - a_3b_2)$, $(a_3b_1 - a_1b_3)$ and $(a_1b_2 - a_2b_1)$ entering into equality (7.55).

It can be easily proved that if the matrix corresponding to a tensor of second rank (L_{ij}) is symmetric (antisymmetric) in one orthonormal basis it is also symmetric (antisymmetric) in any other orthonormal basis.

Finally, every tensor of second rank (L_{ij}) can be represented in the form of a sum of a symmetric tensor and an antisymmetric tensor, namely as

$$L_{ij} = \frac{1}{2}\{L_{ij} + L_{ji}\} + \frac{1}{2}\{L_{ij} - L_{ji}\} \qquad (7.56)$$

where $\frac{L_{ij}+L_{ji}}{2}$ and $\frac{L_{ij}-L_{ji}}{2}$ $(i, j = 1, 2, 3)$ are the components of the symmetric and antisymmetric parts of (L_{ij}) which are uniquely determined by the tensor (L_{ij}).

In § 7 we shall consider an important example of resolution of an orthogonal affine tensor of second rank into the sum of its symmetric and antisymmetric parts, namely we shall resolve the *tensor of relative displacements* into the corresponding symmetric *tensor of pure deformation* and antisymmetric *tensor of rigid body rotation.*

§ 7. TENSOR OF RELATIVE DISPLACEMENTS

Let us consider a state of strain of a physical body (see § 4). Suppose that $\mathbf{U} = \mathbf{U}(\mathbf{r}) = \mathbf{e}_1u_1(x_1, x_2, x_3) + \mathbf{e}_2u_2(x_1, x_2, x_3) + \mathbf{e}_3u_3(x_1, x_2, x_3)$ is the displacement vector of a point specified by the radius vector $\mathbf{r} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$. Under the assumption

that the functions u_1, u_2 and u_3 are differentiable we can write

$$\begin{aligned} du_1 &= \frac{\partial u_1}{\partial x_1}\,dx_1 + \frac{\partial u_1}{\partial x_2}\,dx_2 + \frac{\partial u_1}{\partial x_3}\,dx_3 \\ du_2 &= \frac{\partial u_2}{\partial x_1}\,dx_1 + \frac{\partial u_2}{\partial x_2}\,dx_2 + \frac{\partial u_2}{\partial x_3}\,dx_3 \\ du_3 &= \frac{\partial u_3}{\partial x_1}\,dx_1 + \frac{\partial u_3}{\partial x_2}\,dx_2 + \frac{\partial u_3}{\partial x_3}\,dx_3 \end{aligned} \tag{7.57}$$

Passing from the Cartesian coordinates x_1, x_2, x_3 to new coordinates x_1', x_2', x_3', that is performing a substitution $x_i' = \sum_{k=1}^{3} \alpha_{ik} x_k$ where $\| \alpha_{ik} \|$ is an orthogonal matrix, we can easily verify that the quantities $\frac{\partial u_i}{\partial x_j}$, $i, j = 1, 2, 3$, constitute an orthogonal affine tensor of rank two. This tensor is spoken of as the tensor of relative displacements (corresponding to the state of strain in question). Introducing the notation $\left(\frac{d\mathbf{U}(\mathbf{r})}{d\mathbf{r}}\right)$ for this tensor we can rewrite formulas (7.57) in the form of the equality

$$d\mathbf{U} = \left(\frac{d\mathbf{U}(\mathbf{r})}{d\mathbf{r}}\right) d\mathbf{r} \tag{7.58}$$

Let us now resolve the tensor $\left(\frac{d\mathbf{U}(\mathbf{r})}{d\mathbf{r}}\right)$ into its symmetric and antisymmetric parts. Using the matrix notation we can write down this resolution as

$$\begin{Vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \frac{\partial u_1}{\partial x_3} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \frac{\partial u_2}{\partial x_3} \\ \frac{\partial u_3}{\partial x_1} & \frac{\partial u_3}{\partial x_2} & \frac{\partial u_3}{\partial x_3} \end{Vmatrix} =$$

$$= \begin{Vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{1}{2}\left(\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1}\right) & \frac{1}{2}\left(\frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1}\right) \\ \frac{1}{2}\left(\frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2}\right) & \frac{\partial u_2}{\partial x_2} & \frac{1}{2}\left(\frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2}\right) \\ \frac{1}{2}\left(\frac{\partial u_3}{\partial x_1} + \frac{\partial u_1}{\partial x_3}\right) & \frac{1}{2}\left(\frac{\partial u_3}{\partial x_2} + \frac{\partial u_2}{\partial x_3}\right) & \frac{\partial u_3}{\partial x_3} \end{Vmatrix} +$$

$$+ \begin{Vmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{Vmatrix} \tag{7.59}$$

where ω_1, ω_2 and ω_3 are the coordinates of the vector $\omega = \omega_1 \mathbf{e}_1 + \omega_2 \mathbf{e}_2 + \omega_3 \mathbf{e}_3$ which is equal to $\frac{1}{2}\operatorname{rot} \mathbf{U}$.

The first matrix on the right-hand side of equality (7.59) determines a symmetric tensor D describing a *pure deformation* (without rotation) and the second matrix corresponds to an antisymmetric tensor Ω characterizing a rotation of the body as a whole (i.e. a *rigid body rotation* without deformation). Relation (7.58) can now be put down in the form

$$d\mathbf{U} = \mathrm{D}\, d\mathbf{r} + \Omega\, d\mathbf{r} \tag{7.60}$$

Performing direct calculations we can easily show that

$$\Omega\, d\mathbf{r} = \left[\frac{1}{2} \operatorname{rot} \mathbf{U},\ d\mathbf{r}\right]$$

and consequently

$$d\mathbf{U} = \mathrm{D}\, d\mathbf{r} + \left[\frac{1}{2} \operatorname{rot} \mathbf{U},\ d\mathbf{r}\right] \tag{7.61}$$

Concluding our discussion of a state of strain of a deformable physical body we indicate the following two special cases concerning relative displacements $d\mathbf{U}$ of the points lying in the vicinity of a point $\mathbf{r}$: if the deformation is described by the displacement vectors $\mathbf{U}(\mathbf{r})$ we note that

(1) in case $\operatorname{rot} \mathbf{U} \equiv 0$ it follows from formula (7.61) that the relative displacements $d\mathbf{U}$ are due to a pure deformation;

(2) if $\mathrm{D} = 0$ (i.e. all the elements of the matrix corresponding to the tensor D are equal to zero) the relative displacements $d\mathbf{U}$ are due to a pure rotation.

§ 8. TENSOR FIELD

1. Tensor Field. Divergence of Tensor. If to each point M belonging to a domain G of space there corresponds a tensor (L_{ij}) we say that there is a tensor field (L_{ij})* defined in the domain G. Here the components L_{ij} of the tensor (L_{ij}) are functions of the coordinates of the variable point M (x_1, x_2, x_3).

Characteristic examples of a tensor field are the field of a strain tensor and the field of a stress tensor describing a state of strain and a state of stress of an elastic body subjected to a deformation. Indeed, in the general case a state of strain and a state of stress of such a body vary from point to point and therefore the components of the corresponding tensors depend on the coordinates of the variable point (x_1, x_2, x_3).

Let us suppose that the components L_{ij} of a tensor (L_{ij}) have continuous partial derivatives of the first order with respect to x_1, x_2 and x_3.

* For definiteness, we shall deal with tensor fields constituted by tensors of second rank.

Take the matrix

$$\| L_{ij} \| = \begin{Vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{Vmatrix} \tag{7.62}$$

whose elements are the components of the tensor (L_{ij}) and form the vectors

$$\begin{aligned} \mathbf{L}_1 &= L_{11}\mathbf{e}_1 + L_{12}\mathbf{e}_2 + L_{13}\mathbf{e}_3 \\ \mathbf{L}_2 &= L_{21}\mathbf{e}_1 + L_{22}\mathbf{e}_2 + L_{23}\mathbf{e}_3 \\ \mathbf{L}_3 &= L_{31}\mathbf{e}_1 + L_{32}\mathbf{e}_2 + L_{33}\mathbf{e}_3 \end{aligned} \tag{7.63}$$

The vector $\frac{\partial \mathbf{L}_1}{\partial x_1} + \frac{\partial \mathbf{L}_2}{\partial x_2} + \frac{\partial \mathbf{L}_3}{\partial x_3}$ is called the divergence of the tensor (L_{ij}) and is designated by the symbol div (L_{ij}):

$$\begin{aligned} \operatorname{div}(L_{ij}) &= \frac{\partial \mathbf{L}_1}{\partial x_1} + \frac{\partial \mathbf{L}_2}{\partial x_2} + \frac{\partial \mathbf{L}_3}{\partial x_3} = \left(\frac{\partial L_{11}}{\partial x_1} + \frac{\partial L_{21}}{\partial x_2} + \frac{\partial L_{31}}{\partial x_3} \right) \mathbf{e}_1 + \\ &+ \left(\frac{\partial L_{12}}{\partial x_1} + \frac{\partial L_{22}}{\partial x_2} + \frac{\partial L_{32}}{\partial x_3} \right) \mathbf{e}_2 + \left(\frac{\partial L_{13}}{\partial x_1} + \frac{\partial L_{23}}{\partial x_2} + \frac{\partial L_{33}}{\partial x_3} \right) \mathbf{e}_3 = \\ &= (\operatorname{div}(L_{ij}))_1 \mathbf{e}_1 + (\operatorname{div}(L_{ij}))_2 \mathbf{e}_2 + (\operatorname{div}(L_{ij}))_3 \mathbf{e}_3 \end{aligned} \tag{7.64}$$

The above definition of the divergence of a tensor (L_{ij}) is formal. To justify the definition we must verify whether the divergence thus defined is a vector or, which is the same, whether the quantities

$$\left(\frac{\partial L_{11}}{\partial x_1} + \frac{\partial L_{21}}{\partial x_2} + \frac{\partial L_{31}}{\partial x_3} \right), \quad \left(\frac{\partial L_{12}}{\partial x_1} + \frac{\partial L_{22}}{\partial x_2} + \frac{\partial L_{32}}{\partial x_3} \right) \text{ and } \left(\frac{\partial L_{13}}{\partial x_1} + \frac{\partial L_{23}}{\partial x_2} + \frac{\partial L_{33}}{\partial x_3} \right)$$

constitute a tensor of rank one. Thus, we must prove that the quantities

$$(\operatorname{div}(L_{ij}))_s = \frac{\partial L_{1s}}{\partial x_1} + \frac{\partial L_{2s}}{\partial x_2} + \frac{\partial L_{3s}}{\partial x_3}, \qquad s = 1, 2, 3 \tag{7.65}$$

are transformed like the components of a tensor of rank one when the basis is changed. Let us rewrite expression (7.65) in the form

$$(\operatorname{div}(L_{ij}))_s = \sum_{k=1}^{3} \frac{\partial L_{ks}}{\partial x_k}, \qquad s = 1, 2, 3$$

and pass to a new coordinate system $Ox_1'x_2'x_3'$. In the new system we have

$$(\operatorname{div}(L_{ij}))_p' = \sum_{m=1}^{3} \frac{\partial L_{mp}'}{\partial x_m'} = \sum_{m=1}^{3} \sum_{n=1}^{3} \frac{\partial L_{mp}'}{\partial x_n} \cdot \frac{\partial x_n}{\partial x_m'}, \qquad p = 1, 2, 3 \tag{7.66}$$

The passage from the old Cartesian coordinates to the new ones is performed by means of an orthogonal matrix $\| \alpha_{nm} \|$. But, as is known, the inverse of an orthogonal matrix coincides with its trans

pose, and therefore we have

$$x_n = \sum_{m=1}^{3} \alpha_{mn} x'_m \tag{7.67}$$

According to the definition of a tensor of rank two we can write

$$L'_{mp} = \sum_{k=1}^{3} \sum_{l=1}^{3} \alpha_{mk}\alpha_{pl} L_{kl}, \qquad p = 1, 2, 3 \tag{7.68}$$

Substituting expressions (7.67) and (7.68) into formula (7.66) and taking into account that

$$\sum_{m=1}^{3} \alpha_{mn}\alpha_{mk} = \delta_{nk} = \begin{cases} 1 & \text{for } n = k \\ 0 & \text{for } n \neq k \end{cases}$$

we obtain the relation

$$(\operatorname{div}(L_{ij}))'_p = \sum_{m=1}^{3}\sum_{n=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} \alpha_{mn}\alpha_{mk}\alpha_{pl}\frac{\partial L_{kl}}{\partial x_n} =$$

$$= \sum_{k=1}^{3}\sum_{l=1}^{3} \alpha_{pl}\left[\sum_{n=1}^{3}\left(\sum_{m=1}^{3}\alpha_{mn}\alpha_{mk}\right)\frac{\partial L_{kl}}{\partial x_n}\right] =$$

$$= \sum_{k=1}^{3}\sum_{l=1}^{3} \alpha_{pl}\left(\sum_{n=1}^{3}\delta_{nk}\frac{\partial L_{kl}}{\partial x_n}\right) = \sum_{l=1}^{3}\alpha_{pl}\left(\sum_{k=1}^{3}\frac{\partial L_{kl}}{\partial x_k}\right) =$$

$$= \sum_{l=1}^{3}\alpha_{pl}(\operatorname{div}(L_{ij}))_l, \qquad p = 1, 2, 3$$

which is what we set out to prove.

2. **Ostrogradsky Theorem for Tensor Field.** Let the components L_{ij}, $i = 1, 2, 3$, $j = 1, 2, 3$, of a tensor (L_{ij}) have continuous first-order partial derivatives in a bounded closed domain Ω whose boundary σ_Ω is a piecewise smooth surface. We shall additionally suppose that the domain Ω satisfies the conditions under which the Ostrogradsky theorem for vector functions holds.

Denote the unit outer normal vector to the surface σ_Ω by $\mathbf{n}$ and form, by analogy with relation (7.49′) specifying the product $\mathbf{n}(p_{ij})$, the vector $\mathbf{n}(L_{ij})$. Then we have the formula

$$\iint_{\sigma_\Omega} \mathbf{n}(L_{ij})\,d\sigma = \iiint_{\Omega} \operatorname{div}(L_{ij})\,d\omega \tag{7.69}$$

In other words, the flux of a tensor (L_{ij}) through a closed surface σ_Ω is equal to the triple integral of the divergence of the tensor (L_{ij}) over the volume Ω bounded by the surface.

The flux of a tensor (L_{ij}) through a surface σ_Ω is equal, by definition, to the surface integral on the left-hand side of formula (7.69).

Formula (7.69) expresses the *Ostrogradsky theorem for tensors*.

The proof of formula (7.69) reduces to applying the Ostrogradsky theorem for vectors established in Chapter 5 to each component $L_{1k}\cos(n, x_1)+L_{2k}\cos(n, x_2)+L_{3k}\cos(n, x_3)$ $(k=1, 2, 3)$ of the vector $\mathbf{n}(L_{ij})$:

$$\iint_{\sigma_\Omega}\mathbf{n}(L_{ij})\,d\sigma=\mathbf{e}_1\iint_{\sigma_\Omega}[L_{11}\cos(n, x_1)+L_{21}\cos(n, x_2)+L_{31}\cos(n, x_3)]\,d\sigma+$$

$$+\mathbf{e}_2\iint_{\sigma_\Omega}[L_{12}\cos(n, x_1)+L_{22}\cos(n, x_2)+L_{32}\cos(n, x_3)]\,d\sigma+$$

$$+\mathbf{e}_3\iint_{\sigma_\Omega}[L_{13}\cos(n, x_1)+L_{23}\cos(n, x_2)+L_{33}\cos(n, x_3)]\,d\sigma=$$

$$=\mathbf{e}_1\iiint_{\Omega}(\operatorname{div}(L_{ij}))_1\,d\omega+\mathbf{e}_2\iiint_{\Omega}(\operatorname{div}(L_{ij}))_2\,d\omega+$$

$$+\mathbf{e}_3\iiint_{\Omega}(\operatorname{div}(L_{ij}))_3\,d\omega=\iiint_{\Omega}(\operatorname{div}(L_{ij}))\,d\omega$$

Thus formula (7.69) has been proved.

3. Equations of Motion of a Continuous Medium. Let us apply Ostrogradsky formula (7.69) to deriving an equation of motion of a continuous medium. We mentally isolate an elementary domain

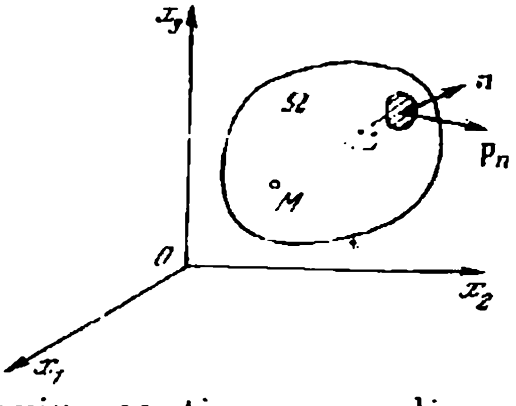


Fig. 7.4

(Ω) occupied by a moving continuous medium (Fig. 7.4). Denoting the surface bounding the domain (Ω) as σ_Ω we can write down the equation expressing Newton's second law for the mass (distributed with density ρ) carried by the domain (Ω). We regard this mass of the moving medium as being concentrated at its centre of gravity, i.e. as a material point. Then the equation can be written in the form

$$\rho\Omega\frac{d\mathbf{v}}{dt}=\rho\Omega\mathbf{f}+\iint_{\sigma_\Omega}\mathbf{p}_n\,d\sigma \tag{7.70}$$

where Ω is the volume of the domain (Ω), $\mathbf{f}$ is the volume force per unit mass and $\mathbf{p}_n$ is the stress on an infinitesimal area $d\sigma$ with unit normal vector $\mathbf{n}$.

The surface integral on the right-hand side of equality (7.70) is equal to

$$\iint_{\sigma_\Omega} \mathbf{p}_n \, d\sigma = \iint_{\sigma_\Omega} \mathbf{n}\Pi \, d\sigma = \iiint_{\Omega} \operatorname{div} \Pi \, d\omega$$

$$= \mathbf{e}_1 \iiint_{\Omega} (\operatorname{div} \Pi)_1 \, d\omega + \mathbf{e}_2 \iiint_{\Omega} (\operatorname{div} \Pi)_2 \, d\omega +$$

$$+ \mathbf{e}_3 \iiint_{\Omega} (\operatorname{div} \Pi)_3 \, d\omega \tag{7.71}$$

where Π is the stress tensor. Applying the mean value theorem to each integral on the right-hand side of equality (7.71) we obtain

$$\iint_{\sigma_\Omega} \mathbf{p}_n \, d\sigma = \mathbf{e}_1\Omega (\operatorname{div} \Pi)_1^* + \mathbf{e}_2\Omega (\operatorname{div} \Pi)_2^* + \mathbf{e}_3\Omega (\operatorname{div} \Pi)_3^* \tag{7.72}$$

where $(\operatorname{div} \Pi)_i^*$ $(i = 1, 2, 3)$ is the value of $(\operatorname{div} \Pi)_i$ assumed at a point $M^* \in (\Omega)$. Now substituting expression (7.72) into (7.70) and passing to the limit as $(\Omega) \to M$ (i.e. as the domain (Ω) is contracted toward an arbitrary fixed point $M \in (\Omega)$ we derive the vector equation of motion of a continuous medium:

$$\rho \frac{d\mathbf{v}}{dt} = \rho \mathbf{f} + \operatorname{div} \Pi \tag{7.73}$$

Vector equation (7.73) can be rewritten in scalar form by projecting its left-hand and right-hand sides on the coordinate axes. This yields the following system of three scalar equations:

$$\begin{aligned} \rho \frac{dv_1}{dt} &= \rho f_1 + \frac{\partial p_{11}}{\partial x_1} + \frac{\partial p_{21}}{\partial x_2} + \frac{\partial p_{31}}{\partial x_3} \\ \rho \frac{dv_2}{dt} &= \rho f_2 + \frac{\partial p_{12}}{\partial x_1} + \frac{\partial p_{22}}{\partial x_2} + \frac{\partial p_{32}}{\partial x_3} \\ \rho \frac{dv_3}{dt} &= \rho f_3 + \frac{\partial p_{13}}{\partial x_1} + \frac{\partial p_{23}}{\partial x_2} + \frac{\partial p_{33}}{\partial x_3} \end{aligned} \tag{7.74}$$

§ 9. PRINCIPAL AXES OF SYMMETRIC TENSOR OF SECOND RANK

Let us take an orthogonal affine tensor (L_{ij}) which we shall interpret as a linear operator (see Sec. 2 in § 2):

$$\mathbf{y} = \mathbf{L}(\mathbf{x}) \tag{7.75}$$

The eigenvectors and eigenvalues of the linear operator $\mathbf{L}(\mathbf{x})$ are referred to as the *eigenvectors* and *eigenvalues of the tensor* (L_{ij}). We remind the reader that an *eigenvector of a linear operator* $\mathbf{L}(\mathbf{x})$ is defined as a nonzero vector $\mathbf{x}$ satisfying the relation

$$\mathbf{L}(\mathbf{x}) = \lambda \mathbf{x} \tag{7.76}$$

where λ is a scalar factor. The number λ is called an *eigenvalue of the operator* L (corresponding to the eigenvector x).

Passing from the vector x to its coordinates x_1, x_2 and x_3 in a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, we can replace vector relation (7.76) by the equivalent system of scalar equalities

$$\left.\begin{aligned} (L_{11}-\lambda)\,x_1 + L_{12}x_2 + L_{13}x_3 = 0 \\ L_{21}x_1 + (L_{22}-\lambda)\,x_2 + L_{23}x_3 = 0 \\ L_{31}x_1 + L_{32}x_2 + (L_{33}-\lambda)\,x_3 = 0 \end{aligned}\right\} \tag{7.77}$$

Let us regard (7.77) as a system of equations in the unknowns x_1, x_2 and x_3. For this system to possess a nontrivial solution, that is one for which x_1, x_2 and x_3 are not simultaneously equal to zero, it is necessary and sufficient that the determinant of system (7.77) turn into zero. Hence, the eigenvalues λ are determined by the equation

$$\begin{vmatrix} L_{11}-\lambda & L_{12} & L_{13} \\ L_{21} & L_{22}-\lambda & L_{23} \\ L_{31} & L_{32} & L_{33}-\lambda \end{vmatrix} = 0 \tag{7.78}$$

As is well known, if (L_{ij}) is a symmetric tensor, i.e. if its matrix is symmetric in every orthonormal basis, all the roots λ_1, λ_2 and λ_3 of equation (7.78) are real. In this case it is possible to construct a system of three unit eigenvectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ associated with the eigenvalues λ_1, λ_2 and λ_3 such that they form an orthonormal basis in which the matrix of the operator L takes the diagonal form

$$\begin{Vmatrix} L_1 & 0 & 0 \\ 0 & L_2 & 0 \\ 0 & 0 & L_3 \end{Vmatrix} \tag{7.79}$$

The vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ thus found specify the so-called principal axes of the tensor (L_{ij}). As an example, we can mention the principal axes of the conductivity tensor of a monocrystal (that is a homogeneous anisotropic body) which are the crystallographic axes. The discussion of the properties of the principal axes of the inertia tensor, strain tensor and stress tensor can be found in courses of theoretical mechanics and mechanics of continua.

§ 10. GENERAL TENSORS

The notion of an orthogonal affine tensor discussed in the foregoing sections is connected with the transformations of orthogonal Cartesian coordinate systems and with the corresponding transformations of their orthonormal bases

Here we shall give the general definition of a tensor which involves all the possible Cartesian coordinate systems (including the oblique ones) specified by the arbitrary bases.

1. Reciprocal Bases. Let

$$\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3 \tag{7.80}$$

be three arbitrary noncomplanar vectors forming a basis in space. For brevity, we shall denote such a basis by a single symbol $\mathbf{e}_i$. Consider the triple scalar product of the vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ which is equal to the volume V of the parallelepiped constructed on these vectors (i.e. the one whose coterminal edges coincide with the vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$):

$$V = (\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3) \tag{7.81}$$

The vectors $\mathbf{e}^k$ $(k = 1, 2, 3)$ determined by the relations

$$\mathbf{e}^1 = \frac{[\mathbf{e}_2,\ \mathbf{e}_3]}{V}, \qquad \mathbf{e}^2 = \frac{[\mathbf{e}_3,\ \mathbf{e}_1]}{V}, \qquad \mathbf{e}^3 = \frac{[\mathbf{e}_1,\ \mathbf{e}_2]}{V} \tag{7.82}$$

constitute a basis which is said to be **reciprocal** to $\mathbf{e}_i$.

We can easily show that, conversely, the basis $\mathbf{e}_i$ is reciprocal to the basis $\mathbf{e}^k$. Indeed, the volume of the parallelepiped whose edges are the vectors $\mathbf{e}^k$ is equal to

$$\begin{aligned} V' &= (\mathbf{e}^1,\ \mathbf{e}^2,\ \mathbf{e}^3) = \left(\frac{[\mathbf{e}_2,\ \mathbf{e}_3]}{V},\ \frac{[\mathbf{e}_3,\ \mathbf{e}_1]}{V},\ \frac{[\mathbf{e}_1,\ \mathbf{e}_2]}{V}\right) = \\ &= \frac{1}{V^3}\left([\mathbf{e}_2,\ \mathbf{e}_3]\left[[\mathbf{e}_3,\ \mathbf{e}_1],\ [\mathbf{e}_1,\ \mathbf{e}_2]\right]\right) = \\ &= \frac{1}{V^3}\left([\mathbf{e}_2,\ \mathbf{e}_3]\ \left\{\mathbf{e}_1(\mathbf{e}_3\mathbf{e}_1\mathbf{e}_2) - \mathbf{e}_2(\mathbf{e}_1\mathbf{e}_1\mathbf{e}_3)\right\}\right) = \frac{V^2}{V^3} = \frac{1}{V} \end{aligned} \tag{7.83}$$

Therefore we have

$$VV' = 1 \tag{7.84}$$

thus

$$\frac{[\mathbf{e}^2,\ \mathbf{e}^3]}{V'} = V\left[\frac{[\mathbf{e}_3,\ \mathbf{e}_1],\ [\mathbf{e}_1,\ \mathbf{e}_2]}{V^2}\right] = \frac{V^2\mathbf{e}_1}{V^2} = \mathbf{e}_1 \tag{7.85}$$

and, similarly,

$$\left[\frac{\mathbf{e}^3,\ \mathbf{e}^1}{V'}\right] = \mathbf{e}_2, \qquad \left[\frac{\mathbf{e}^1,\ \mathbf{e}^2}{V'}\right] = \mathbf{e}_3 \tag{7.86}$$

From relations (7.85) and (7.86) it follows that

$$(\mathbf{e}_i,\ \mathbf{e}^k) = \delta_i^k = \begin{cases} 0 & \text{for } i \neq k \\ 1 & \text{for } i = k \end{cases} \tag{7.87}$$

It should be noted that every orthonormal basis e_1, e_2, e_3 coincides with its reciprocal basis.

2. **Covariant and Contravariant Components of Vector.** Let us take a basis e_i and its reciprocal basis e^i and write down the resolutions of an arbitrary vector x with respect to the bases:

$$x = \sum_{i=1}^{3} x_i e^i = \sum_{i=1}^{3} x^i e_i \tag{7.88}$$

The coefficients entering into the resolution of the vector x relative to the given basis are called the **contravariant components (contravariant coordinates)** of the vector x in this basis. Thus, the numbers x^i and x_i are, respectively, the contravariant components of the vector x in the bases e_1, e_2, e_3 and e^1, e^2, e^3.

The **covariant components** of a vector x in a given basis are the scalar products of the vector by the base vectors of the reciprocal basis.

Multiplying scalarly equalities (7.88) by e_k (e^k) and taking advantage of relations (7.87) we find that the covariant components of the vector x in the bases e^1, e^2, e^3 (e_1, e_2, e_3) are respectively equal to

$$(x, e_k) = \sum_{i=1}^{3} x_i (e^i, e_k) = \sum_{i=1}^{3} x_i \delta_k^i = x_k \tag{7.89}$$

and

$$(x, e^k) = \sum_{i=1}^{3} x^i (e_i, e^k) = \sum_{i=1}^{3} x^i \delta_i^k = x^k \tag{7.90}$$

Consequently, the covariant components of a vector in a given basis are its contravariant components in the reciprocal basis.

3. **Summation Convention.** In the theory of tensors we usually follow a *summation convention* (due to A. Einstein*) which applies as follows: if a subscript and a superscript entering into an expression are labelled by the same symbol, this symbol (index) is understood as denoting a summation with respect to that index over its range.

In what follows the indices under consideration take on the values 1, 2 and 3 and hence a summation, if necessary, is carried out with respect to an index ranging from one to three. For instance, applying this rule to the sums entering into formula (7.88) we can write down the resolutions of the vector x in the form

$$x = x_i e^i, \quad x = x^i e_i \tag{7.91}$$

* Einstein, Albert (1879-1955), the great 20th century physicist, the creator of the theory of relativity (born in Germany).

Similary, a bilinear form $\sum_{i,\,k=1}^{3} a_{ik}x^ix^k$ is written as

$$a_{ik}x^ix^k \tag{7.92}$$

etc.

4. Transformation of Base Vectors. Let us consider the transformation from an old basis $\mathbf{e}_i$ to a new basis $\mathbf{e}_{i'}$. Using the summation convention we can write

$$\mathbf{e}_{i'} = \alpha^i_{i'}\mathbf{e}_i, \quad i' = 1,\ 2,\ 3 \tag{7.93}$$

where the coefficients $\alpha^i_{i'}$ form the transformation matrix $\|\alpha^i_{i'}\|$ from the old basis $\mathbf{e}_i$ to the new basis $\mathbf{e}_{i'}$, i.e.

$$\|\alpha^i_{i'}\| = \left\| \begin{matrix} \alpha^1_{1'} & \alpha^2_{1'} & \alpha^3_{1'} \\ \alpha^1_{2'} & \alpha^2_{2'} & \alpha^3_{2'} \\ \alpha^1_{3'} & \alpha^2_{3'} & \alpha^3_{3'} \end{matrix} \right\| \tag{7.94}$$

If we consider the inverse transformation from the new basis $\mathbf{e}_{i'}$ to the old basis $\mathbf{e}_i$ which is written as

$$\mathbf{e}_i = \alpha^{i'}_i\mathbf{e}_{i'} \tag{7.95}$$

the matrix $\|\alpha^{i'}_i\|$ is obviously the inverse of the matrix $\|\alpha^i_{i'}\|$. Actually, substituting the expression $\mathbf{e}_{i'} = \alpha^i_{i'}\mathbf{e}_i$ into the equality

$$\mathbf{e}_k = \alpha^{i'}_k\mathbf{e}_{i'} \tag{7.96}$$

we obtain

$$\mathbf{e}_k = \alpha^{i'}_k\alpha^i_{i'}\mathbf{e}_i \tag{7.97}$$

The resolution of each vector $\mathbf{e}_k$ $(k = 1,\ 2,\ 3)$ with respect to the base vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ being unique, we derive from formulas (7.97) the relations

$$\alpha^{i'}_k\alpha^i_{i'} - \delta^i_k = \begin{cases} 0 & \text{for } i \neq k \\ 1 & \text{for } i = k \end{cases} \tag{7.98}$$

which suggest that the matrices $\|\alpha^{i'}_i\|$ and $\|\alpha^i_{i'}\|$ are mutually inverse.

5. Transformation of Covariant and Contravariant Components of Vector. Let us first consider the transformation of the contravariant components of an arbitrary vector

$$\mathbf{x} = x^i\mathbf{e}_i = x^{i'}\mathbf{e}_{i'} \tag{7.99}$$

when a basis $\mathbf{e}_i$ is transformed to a basis $\mathbf{e}_{i'}$. Substituting the expression $\mathbf{e}_i = \alpha^{i'}_i\mathbf{e}_{i'}$ into formula (7.99) we see that

$$\mathbf{x} = x^i\alpha^{i'}_i\mathbf{e}_{i'} = x^{i'}\mathbf{e}_{i'}$$

By the uniqueness of the resolution of the vector $\mathbf{x}$ in the basis $\mathbf{e}_{1'}$, $\mathbf{e}_{2'}$, $\mathbf{e}_{3'}$, we can write

$$x^{i'} = \alpha_i^{i'} x^i \tag{7.100}$$

Thus, the "new" contravariant coordinates $x^{i'}$ are expressed in terms of the "old" contravariant coordinates x^i by means of the matrix $\| \alpha_i^{i'} \|$ specifying the inverse transformation from the new basis $\mathbf{e}_{i'}$ to the old basis $\mathbf{e}_i$*. This accounts for the term "contravariant components" which indicates that the expressions of $\mathbf{e}_i$ in terms of $\mathbf{e}_{i'}$ and of $x^{i'}$ in terms of x^i involve, respectively, the elements of the matrix $\|\alpha_i^{i'}\|$ and of its inverse.

By analogy with (7.100), we obtain the relation

$$x^i = \alpha_{i'}^i x^{i'} \tag{7.101}$$

Let us now proceed to investigate the transformation of the covariant components x_i of a vector $\mathbf{x}$. We have

$$x_i = (\mathbf{x}, \mathbf{e}_i), \qquad x_{i'} = (\mathbf{x}, \mathbf{e}_{i'}) \tag{7.102}$$

and consequently

$$x_{i'} = (\mathbf{x}, \mathbf{e}_{i'}) = (\mathbf{x}, \alpha_{i'}^i \mathbf{e}_i) = \alpha_{i'}^i x_i \tag{7.103}$$

Similarly, by analogy with (7.103) we obtain

$$x_i = \alpha_i^{i'} x_{i'} \tag{7.104}$$

Thus, the transformation of the covariant components of a vector is performed by means of the same matrix as the transformation of the base vectors. The terms "covariant components" indicated this coincidence of the matrices.

6. **General Definition of Tensor.** As before, we shall denote the transformation matrix from an old basis $\mathbf{e}_i$ to a new basis $\mathbf{e}_{i'}$ by $\| \alpha_{i'}^i \|$ and the matrix of the inverse transformation from the new basis $\mathbf{e}_{i'}$ to the old basis $\mathbf{e}_i$ by $\| \alpha_i^{i'} \|$.

Definition 1. *A quantity A which is specified in every basis $\mathbf{e}_i$ ($i = 1, 2, 3$) by means of 3^{p+q} numbers $A_{i_1 i_2 \ldots i_p}^{j_1 j_2 \ldots j_q}$ where the indices i_s, $s = 1, 2, 3, \ldots, p$, and j_t, $t = 1, 2, \ldots, q$, independently assume the values 1, 2 and 3 is called a **tensor of rank (order)** $p+q$ (**p-fold covariant** and **q-fold contravariant**) if these numbers undergo the transformation determined by the formulas*

$$A_{i'_1 i'_2 \ldots i'_p}^{j'_1 j'_2 \ldots j'_q} = \alpha_{i'_1}^{i_1} \alpha_{i'_2}^{i_2} \ldots \alpha_{i'_p}^{i_p} \alpha_{j_1}^{j'_1} \alpha_{j_2}^{j'_2} \ldots \alpha_{j_q}^{j'_q} A_{i_1 i_2 \ldots i_p}^{j_1 j_2 \ldots j_q} \tag{7.105}$$

---

* More precisely, formulas (7.100) show that the coefficients entering into the expressions of $x^{i'}$ in terms of x^i constitute a matrix which is the transpose of the inverse matrix $\| \alpha_{i'}^i \|$.

when we pass from an arbitrary basis ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$) *to any other basis* $\mathbf{e}_{1'}$, $\mathbf{e}_{2'}$, $\mathbf{e}_{3'}$ where $\| \alpha^{i}_{i'} \|$ *is the transformation matrix from the basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *to the basis* $\mathbf{e}_{1'}$, $\mathbf{e}_{2'}$, $\mathbf{e}_{3'}$ *and* $\| \alpha^{i'}_{i} \|$ *is its inverse matrix.*

The numbers $A^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ are spoken of as the components of the tensor A relative to the basis $\mathbf{e}_i$. The superscripts $j_1, \ldots, j_q$ are called the contravariant indices of the tensor and the subscripts $i_1, \ldots, i_p$ are its covariant indices.

There is an alternative form of definition of a tensor (equivalent to the above):

Let, in every basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, *there be given a system of* 3^{p+q} *numbers* $A^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ *where the indices* i_s, $s = 1, 2, \ldots, p$, *and* j_t, $t = 1, 2, \ldots, q$, *independently assume the values* 1, 2 *and* 3. *If the passage to any other basis* $\mathbf{e}_{1'}$, $\mathbf{e}_{2'}$, $\mathbf{e}_{3'}$ *results in the transformation of these numbers according to formulas* (7.105):

$$A^{j'_1 j'_2 \ldots j'_q}_{i'_1 i'_2 \ldots i'_p} = \alpha^{i_1}_{i'_1} \alpha^{i_2}_{i'_2} \ldots \alpha^{i_p}_{i'_p} \alpha^{j'_1}_{j_1} \alpha^{j'_2}_{j_2} \ldots \alpha^{j'_q}_{j_q} A^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$$

where $\| \alpha^{i}_{i'} \|$ *is the transformation matrix from the basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *to the basis* $\mathbf{e}_{1'}$, $\mathbf{e}_{2'}$, $\mathbf{e}_{3'}$ *and* $\| \alpha^{i'}_{i} \|$ *is the inverse of* $\| \alpha^{i}_{i'} \|$ *we say that we are given a tensor of rank* $p+q$, *p-fold covariant and q-fold contravariant* (*or* ***covariant of rank*** *p* ***and contravariant of rank*** *q*).

Examples

1. (a) The system of coefficients $a_{i_1 i_2}$ of an invariant bilinear form

$$f(\mathbf{x}, \mathbf{y}) = a_{i_1 i_2} x^{i_1} y^{i_2} = a_{i'_1 i'_2} x^{i'_1} y^{i'_2}, \tag{7.106}$$

$$(\mathbf{x} = x^{i_1}\mathbf{e}_{i_1} = x^{i'_1}\mathbf{e}_{i'_1}, \ \mathbf{y} = y^{i_2}\mathbf{e}_{i_2} = y^{i'_2}\mathbf{e}_{i'_2},$$
$$i_1 = 1, 2, 3; \ i_2 = 1, 2, 3; \ i'_1 = 1, 2, 3; \ i'_2 = 1, 2, 3)$$

is a *covariant tensor* of second rank, i.e. having only covariant indices.

Indeed, substituting the expressions

$$x^{i_1} = \alpha^{i_1}_{i'_1} x^{i'_1} \quad \text{and} \quad y^{i_2} = \alpha^{i_2}_{i'_2} y^{i'_2} \tag{7.107}$$

into formula (7.106) we obtain an identity involving the coordinates of two arbitrary vectors $\mathbf{x}$ and $\mathbf{y}$, which implies that

$$a_{i'_1 i'_2} = \alpha^{i_1}_{i'_1} \alpha^{i_2}_{i'_2} a_{i_1 i_2} \tag{7.108}$$

(b) In particular, if $f(\mathbf{x}, \mathbf{y})$ is equal to the scalar product $(\mathbf{x}, \mathbf{y})$ of two vectors $\mathbf{x}$ and $\mathbf{y}$, the collection of the coefficients $g_{i_1 i_2}$ of the bilinear form

$$(\mathbf{x}, \mathbf{y}) = g_{i_1 i_2} x^{i_1} y^{i_2}$$

is termed a (fundamental) metric tensor or a covariant metric tensor. The elements of the inverse of the matrix $\|g_{i_1 i_2}\|$ (denoted by the symbols $g^{i_1 i_2}$) form a so-called contravariant metric tensor.

By the symmetry property of scalar product ((x, y) = (y, x)), the tensors g_{ij} and g^{ij} are symmetric, that is we have, in every basis, the relations

$$g_{ij} = g_{ji}, \quad g^{ij} = g^{ji}$$

2. The elements L_i^j of the matrix of a linear operator L determined by the relations

$$\mathrm{L}(\mathbf{e}_i) = L_i^j \mathbf{e}_j, \quad i = 1, 2, 3 \tag{7.109}$$

constitute a tensor (of second order) covariant of rank one and contravariant of rank one. For, in the new basis $\mathbf{e}_{i'}$, we have

$$\mathrm{L}(\mathbf{e}_{i'}) = L_{i'}^{j'} \mathbf{e}_{j'} \tag{7.110}$$

On the other hand, we have

$$\mathrm{L}(\mathbf{e}_{i'}) = \mathrm{L}(\alpha_{i'}^{i} \mathbf{e}_i) = \alpha_{i'}^{i} \mathrm{L}(\mathbf{e}_i) = \alpha_{i'}^{i} L_i^j \alpha_j^{j'} \mathbf{e}_{j'} \tag{7.111}$$

Comparing (7.110) with (7.111), by the uniqueness of resolution of every vector $\mathrm{L}(\mathbf{e}_i)$ with respect to the basis $\mathbf{e}_{j'}$, we obtain

$$L_{i'}^{j'} = \alpha_{i'}^{i} \alpha_j^{j'} L_i^j \tag{7.112}$$

3. The covariant components x_i of a vector x constitute a covariant tensor of rank one, and the contravariant coordinates x^i of x form a contravariant tensor of rank one.

7. Operations on Tensors. In the general case the operations on tensors are defined in the same manner as for the orthogonal affine tensors. The operations of addition and subtraction are naturally defined only for the tensors of the same rank having the same number p of covariant indices and the same number q of contravariant indices. The contraction is applied only to mixed tensors (i.e. having both contravariant and covariant indices) by putting one contravariant index equal to a covariant index and summing with respect to that index.

There are also some other operations such as *raising* or *lowering* indices by means of multiplying the fundamental (covariant) metric tensor or the contravariant metric tensor by a given tensor T and then contracting the product by putting a covariant (contravariant) index of g_{ij} (g^{ij}) equal to a contravariant (covariant) index of T and summing with respect to that index etc.

8. Some Further Generalizations. Further generalizations are connected with the introduction of curvilinear coordinates. This gives rise to some new notions such as covariant and contravariant differentiation of a tensor and others. For the general theory of tensors we refer the reader to [4], [10] and [14].

APPENDIX TO CHAPTER 7

ON MULTIPLICATION OF MATRICES

We remind the reader that the product $P \cdot Q$ of two rectangular matrices P and Q is only defined for the case when the number of columns in the first factor equals the number of rows in the second factor. If

$$P = \begin{Vmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ . & . & . & . \\ p_{m1} & p_{m2} & \dots & p_{mn} \end{Vmatrix} \quad \text{and} \quad Q = \begin{Vmatrix} q_{11} & q_{12} & \dots & q_{1s} \\ q_{21} & q_{22} & \dots & q_{2s} \\ . & . & . & . \\ q_{n1} & q_{n2} & \dots & q_{ns} \end{Vmatrix}$$

are such matrices their *product* is the matrix

$$R = \begin{Vmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ . & . & . & . \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{Vmatrix}$$

whose element r_{ij} ($i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$) belonging to the ith row and the jth column is determined by the formula

$$r_{ij} = \sum_{\nu=1}^{n} p_{i\nu} q_{\nu j}$$

Hence, if we interpret the elements of the ith row of P as the coordinates of an n-dimensional vector and the elements of the jth column of Q as the coordinates of another n-dimensional vector we can say that r_{ij} is equal to the scalar product of the ith row of the first factor by the jth column of the second factor.

Let us take two vectors $\mathbf{x}$ and $\mathbf{y}$ and represent them as the column matrices

$$\mathbf{x} = \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} \quad \text{and} \quad \mathbf{y} = \begin{Vmatrix} y_1 \\ y_2 \\ y_3 \end{Vmatrix}$$

where x_1, x_2, x_3 and y_1, y_2, y_3 are the coordinates of $\mathbf{x}$ and $\mathbf{y}$ in a given basis. If $\| L_{ij} \|$ is a matrix of the form

$$\mathrm{L} = \begin{Vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{Vmatrix}$$

then, by definition, the equality $\mathbf{y} = \mathrm{L}\mathbf{x}$ is equivalent to the relation

$$\begin{Vmatrix} y_1 \\ y_2 \\ y_3 \end{Vmatrix} = \begin{Vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} \tag{1}$$

After the multiplication of the matrices on the right-hand side of (1) has been performed we obtain

$$\begin{Vmatrix} y_1 \\ y_2 \\ y_3 \end{Vmatrix} = \begin{Vmatrix} \sum\limits_{k=1}^{3} L_{1k}x_k \\ \sum\limits_{k=1}^{3} L_{2k}x_k \\ \sum\limits_{k=1}^{3} L_{3k}x_k \end{Vmatrix} \tag{2}$$

which is equivalent to the three scalar equalities

$$\left.\begin{aligned} y_1 &= L_{11}x_1 + L_{12}x_2 + L_{13}x_3 \\ y_2 &= L_{21}x_1 + L_{22}x_2 + L_{23}x_3 \\ y_3 &= L_{31}x_1 + L_{32}x_2 + L_{33}x_3 \end{aligned}\right\} \tag{3}$$

As has been said, relation (1) (or equivalent relations (2) and (3)) is the definition of the *multiplication* (*on the right*) of a matrix by a vector.

We can similarly take two vectors $\mathbf{x}$ and $\mathbf{y}^*$ and represent them as the row matrices

$$\mathbf{x} = \| x_1, x_2, x_3 \| \quad \text{and} \quad \mathbf{y}^* = \| y_1^*, y_2^*, y_3^* \|$$

Then an equality of the form

$$\| y_1^*, y_2^*, y_3^* \| = \| x_1, x_2, x_3 \| \cdot \begin{Vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{Vmatrix} \tag{4}$$

can be rewritten as

$$\| y_1^*, y_2^*, y_3^* \| = \left\| \sum_{i=1}^{3} L_{i1}x_i, \quad \sum_{i=1}^{3} L_{i2}x_i, \quad \sum_{i=1}^{3} L_{i3}x_i \right\|$$

after the multiplication of the matrices on the right-hand side has been performed. The latter relation is equivalent to the three scalar equalities

$$\left.\begin{aligned} y_1^* &= L_{11}x_1 + L_{21}x_2 + L_{31}x_3 \\ y_2^* &= L_{12}x_1 + L_{22}x_2 + L_{32}x_3 \\ y_3^* &= L_{13}x_1 + L_{23}x_2 + L_{33}x_3 \end{aligned}\right\} \tag{5}$$

In the contracted notation equality (1) is written in the form

$$\mathbf{y} = \mathbf{L}\mathbf{x} \tag{6}$$

and equality (4) in the form

$$\mathbf{y}^* = \mathbf{x}\mathbf{L} \tag{7}$$

where $\mathbf{L}$ is *multiplied by* $\mathbf{x}$ *on the left*.

8 Functional Sequences and Series

In this chapter we shall study sequences and series (referred to as functional sequences and series) whose members are functions.

In practical applications we often try to expand a given function in a functional series whose terms are functions which are in a certain sense simpler than the given function. Such an expansion facilitates the investigation of the function, the computation of its values and the integration. Functional series are also used in the theory of differential equations and other divisions of mathematics and its applications.

In investigating the properties of functional series and sequences we introduce various types of convergence. Among them, *uniform convergence* and *convergence in the mean* are of particular importance.

§ 1. UNIFORM CONVERGENCE. TESTS FOR UNIFORM CONVERGENCE

1. Convergence and Uniform Convergence. Let us consider a sequence of functions

$$f_1(x),\ f_2(x),\ \ldots,\ f_n(x),\ \ldots \tag{8.1}$$

defined on a closed interval $a \leqslant x \leqslant b$.* If an arbitrary fixed value $x_0 \in [a, b]$ is substituted for the current variable x functional sequence (8.1) turns into a numerical sequence of the form

$$f_1(x_0),\ f_2(x_0),\ \ldots,\ f_n(x_0),\ \ldots \tag{8.2}$$

Functional sequence (8.1) is said to be convergent at a point x_0 if number sequence (8.2) is convergent. Functional sequence (8.1) is said to be divergent at a point x_0 if sequence (8.2) is divergent.

---

* Instead of a closed interval $a \leqslant x \leqslant b$ we can take any other set X of values of x, for instance, $a < x < b$, $a \leqslant x < b$, $a < x \leqslant b$, $a < x < +\infty$, $a \leqslant x < +\infty$, $-\infty < x < +\infty$ etc. In what follows we shall stipulate the cases when such a replacement of a closed interval $[a, b]$ by an arbitrary set is inadmissible.

Accordingly, in the former case x_0 is called a point of convergence of sequence (8.1) and in the latter case a point of divergence.*

If a functional sequence converges at each point $x \in [a, b]$ we say that it converges on the interval $[a, b]$. If a sequence $\{f_n(x)\}$, $n = 1, 2, \ldots$, converges on $[a, b]$ there exists a certain limit $\lim\limits_{n\to\infty} f_n(x)$ (which in the general case can vary from point to point) at each point x of the interval $[a, b]$. Therefore this limit is a function $f(x)$ defined on $[a, b]$. The function $f(x)$ is called the limit of functional sequence (8.1) and we write

$$f_n(x) \to f(x) \quad \text{as} \quad n \to +\infty$$

or

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{on} \quad [a, b] \tag{8.3}$$

We can now formulate the following

Definition 1. *A functional sequence $\{f_n(x)\}$ is said to be* ***convergent*** *to a function $f(x)$ on the interval $[a, b]$ if for each fixed value $x \in [a, b]$ the number sequence $f_n(x)$, $n = 1, 2, \ldots$, converges to the number $f(x)$, that is if for every $\varepsilon > 0$ and every $x \in [a, b]$ there is a number $N = N(\varepsilon, x)$*** *(dependent on ε and, generally speaking, on x) such that*

$$|f_n(x) - f(x)| < \varepsilon \quad \textit{for every} \quad n > N(\varepsilon, x) \tag{8.4}$$

Among the convergent functional sequences the so-called *uniformly convergent* sequences are essentially important.

Definition 2. *A functional sequence $\{f_n(x)\}$ is called* ***uniformly convergent*** *on an interval $[a, b]$ to a function $f(x)$ if, given any $\varepsilon > 0$, there exists a number $N = N(\varepsilon)$**** *(dependent on ε but independent of x) such that the difference between $f_n(x)$ and $f(x)$ satisfies the condition*

$$|f_n(x) - f(x)| < \varepsilon \quad \textit{for every} \quad n > N(\varepsilon) \tag{8.5}$$

for all $x \in [a, b]$ simultaneously.

This definition can be restated in an equivalent form:

---

* The set of all the points of convergence of functional sequence (8.1) is referred to as the domain (or region) of convergence of the sequence. The domain of convergence of a functional sequence can be an arbitrary set of any complex structure. It may coincide with the whole x-axis (as in the case of the sequence $f_n(x) \equiv \dfrac{1}{n}$, $-\infty < x < +\infty$, $n = 1, 2, \ldots$, convergent on the entire x-axis to the function $f(x) \equiv 0$) or be an empty set containing no points (e.g. for the sequence $f_n(x) = (-1)^n$, $-\infty < x < +\infty$, $n = 1, 2, \ldots$, which diverges at every point $x \in (-\infty, +\infty)$).

** The number $N(\varepsilon, x)$ may not be an integer.

*** The number $N(\varepsilon)$ is not necessarily an integer.

Definition 2'. *A functional sequence $\{f_n(x)\}$ is said to converge uniformly to a function $f(x)$ on an interval $[a, b]$ if*

$$\sup_{a \leqslant x \leqslant b} |f_n(x) - f(x)| \to 0 \quad \textit{for} \quad n \to +\infty \tag{8.5'}$$

*that is if the least upper bound of $|f_n(x) - f(x)|$ (spoken of as the **maximum deviation of the function** $f_n(x)$ **from the function** $f(x)$ **on the interval** $[a, b]$) tends to zero as $n \to \infty$.*

Indeed, if condition (8.5') is fulfilled then for every $\varepsilon > 0$ there is $N(\varepsilon)$ such that for any $n > N(\varepsilon)$ the inequality

$$\sup_{a \leqslant x \leqslant b} |f_n(x) - f(x)| < \varepsilon$$

holds for all $x \in [a, b]$. But, by the definition of the least upper bound, we have

$$|f_n(x) - f(x)| \leqslant \sup_{a \leqslant x \leqslant b} |f_n(x) - f(x)|$$

for all $x \in [a, b]$. Therefore relations (8.5) are also fulfilled.

Conversely, if relations (8.5) take place we have

$$\sup_{a \leqslant x \leqslant b} |f_n(x) - f(x)| < \varepsilon$$

for every $n > N(\varepsilon)$, which implies (8.5') since $\varepsilon > 0$ has been chosen quite arbitrarily.

Uniform convergence of a sequence $\{f_n(x)\}$ to a function $f(x)$ on $[a, b]$ will be designated by the symbol relation

$$f_n(x) \rightrightarrows f(x) \quad \text{on} \quad [a, b] \tag{8.6}$$

The notion of uniform convergence admits of a simple geometric interpretation. Relation (8.5') means that the least upper bound of the deviation of the graph of the function $y = f_n(x)$ from that

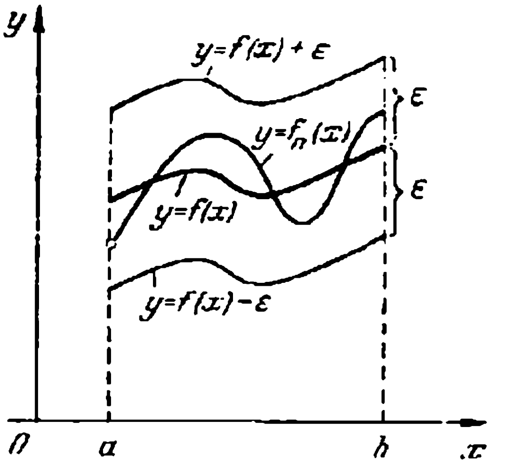


Fig. 8.1

of the function $y = f(x)$ on the interval $[a, b]$ tends to zero as $n \to +\infty$. In other words, if we envelope the graph of the function $y = f(x)$ by an "ε-strip" (shown in Fig. 8.1) determined by the relations

$$f(x) - \varepsilon < y < f(x) + \varepsilon, \qquad a \leqslant x \leqslant b \tag{8.7}$$

then, beginning with a sufficiently large n, the graphs of all the functions $y = f_n(x)$ entirely lie within the ε-strip enveloping the graph of the limit function $f(x)$.

Examples

1. The sequence $f_n(x) = \frac{1}{n} \sin nx$ converges to $f(x) \equiv 0$ as $n \to +\infty$ on the entire x-axis $-\infty < x < +\infty$. Here the convergence is uniform because $| f_n(x) - f(x) | = \frac{1}{n} | \sin nx | \leqslant \frac{1}{n} < \varepsilon$ for all x, $-\infty < x < +\infty$, simultaneously provided $n > N(\varepsilon) = \frac{1}{\varepsilon}$.

2. The sequence $f_n(x) = x^n$ converges, as $n \to +\infty$, on the interval $0 \leqslant x \leqslant 1$ to the function $f(x)$ determined by the relations

$$f(x) = \begin{cases} 0 & \text{for } 0 \leqslant x < 1 \\ 1 & \text{for } x = 1 \end{cases}$$

But here the convergence is nonuniform. For, if we take $0 < \varepsilon < 1$ and $0 < x < 1$ the inequality $| f_n(x) - f(x) | = x^n < \varepsilon$ holds only when $n > N(\varepsilon, x) = \frac{\ln \varepsilon}{\ln x}$, and $N(\varepsilon, x) = \frac{\ln \varepsilon}{\ln x} \to +\infty$ if $x \to 1 - 0$ for every fixed $\varepsilon \in (0, 1)$. Consequently, for every ε taken from the interval $0 < \varepsilon < 1$ there is no finite $N(\varepsilon)$ independent of x such that the inequality $| f_n(x) - f(x) | = x^n < \varepsilon$

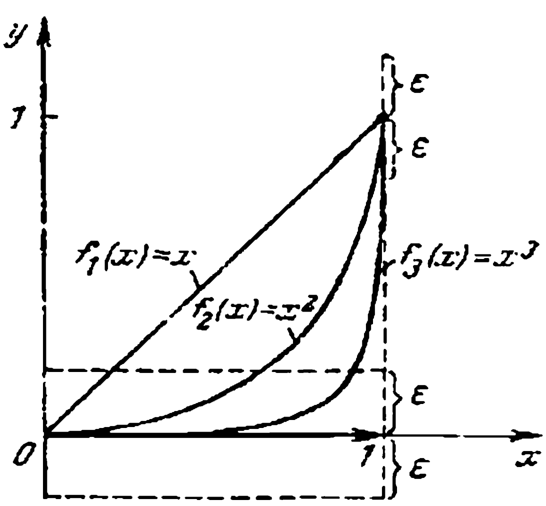


Fig. 8.2

holds for every $n > N(\varepsilon)$ and for all x belonging to the half-open interval $0 \leqslant x < 1$. If we replace the segment $0 \leqslant x \leqslant 1$ by a smaller segment $0 \leqslant x \leqslant 1 - \delta$ with an arbitrarily small δ, $0 < \delta < 1$, the sequence $f_n(x) = x^n$ converges uniformly to its limit $f(x) \equiv 0$ on this smaller interval. Indeed, we have

$$N(\varepsilon, x) = \frac{\ln \varepsilon}{\ln x} \leqslant N(\varepsilon) = \frac{\ln \varepsilon}{\ln(1-\delta)} \quad \text{for } 0 \leqslant x \leqslant 1 - \delta$$

and therefore $|f_n(x)-f(x)|=x^n<\varepsilon$ for all $x\in[0, 1-\delta]$ when $n>N(\varepsilon)=\frac{\ln\varepsilon}{\ln(1-\delta)}$.

From the geometrical point of view this example can be interpreted as follows. In Fig. 8.2 we see the graphs of several functions $f_n(x)$ belonging to the sequence and the graph of the limiting function $f(x)$, the latter being shown in the heavy line. The graph of $f(x)$ consists of the half-segment $0\leqslant x<1$ (with the end point $x=1$ excluded) of the x-axis and an isolated point with the coordinates $(1, 1)$. Let us envelope the graph of the limit function by an "ε-strip", $0<\varepsilon<1$. The graph of every function $f_n(x)=x^n$ starts from the origin of coordinates and its right end point lies at the point $(1, 1)$. Therefore the function $f_n(x)=x^n$ being continuous, its graph must leave the "ε-strip" at a point x, $0<x<1$. Hence, the sequence $f_n(x)=x^n$, $n=1, 2, \ldots$, converges nonuniformly on the interval $0\leqslant x\leqslant 1$.

3. The functional sequence $f_n(x)=\frac{2}{\pi}\arctan nx$, $-\infty<x<+\infty$, $n=1, 2, 3, \ldots$, converges to the function

$$f(x)=\operatorname{sgn} x=\begin{cases}-1 & \text{for } -\infty<x<0\\ 0 & \text{for } x=0\\ +1 & \text{for } 0<x<+\infty\end{cases}$$

but the sequence does not converge uniformly which can be easily established by the geometric method applied in the foregoing example.

4. The sequence of functions $f_n(x)=\frac{2nx}{1+n^2x^2}$, $n=1, 2, \ldots$, converges to the function $f(x)=0$ on the positive x axis $0\leqslant x<+\infty$. To find out whether the sequence is uniformly convergent to its limit on the half-line $0\leqslant x<+\infty$ we shall check up the validity of relation (8.5′). Thus, we must verify if $\sup\limits_{0\leqslant x<+\infty}|f_n(x)-f(x)|\to 0$ for $n\to+\infty$. To this end we evaluate the maximum of $\varphi_n(x)=|f_n(x)-f(x)|=\frac{2nx}{1+n^2x^2}$ on the positive half of x-axis. We have

$$\varphi_n'(x)=\frac{(1+n^2x^2)2n-2nx\cdot 2n^2x}{(1+n^2x^2)^2}=2n\frac{1-n^2x^2}{(1+n^2x^2)^2}$$

and hence $\varphi_n'(x)=0$ for $1-n^2x^2=0$, i.e. for $x_n=\frac{1}{n}$. Consequently,

$$\max_{0\leqslant x<+\infty}\varphi_n(x)=\varphi_n\left(\frac{1}{n}\right)=\frac{2n\cdot\frac{1}{n}}{1+n^2\frac{1}{n^2}}=1\not\to 0 \text{ for } n\to\infty$$

Therefore the sequence does not converge uniformly. In this example nonuniform convergence is due to the fact that the maximum value of $f_n(x)$ which is equal to unity coincides with the maximum deviation of the graph of $f_n(x)$ from the graph of $f(x)$ on the interval $0 \leqslant x < +\infty$, the point of maximum $x_n = \frac{1}{n}$ moving to the left as $n \to \infty$ (see Fig. 8.3).

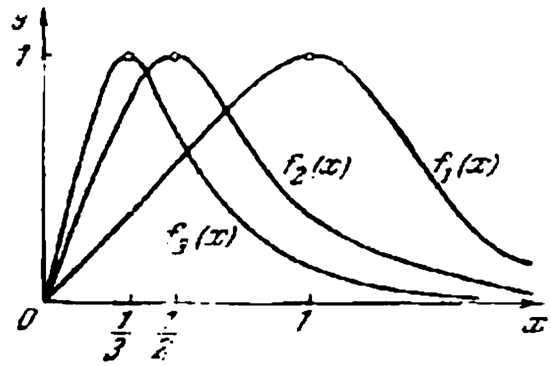


Fig. 8.3

The notions introduced for the functional sequences can be easily transferred to the functional series of the form

$$\sum_{k=1}^{\infty} u_k(x) = u_1(x) + u_2(x) + \ldots + u_k(x) + \ldots \tag{8.8}$$

where the functions $u_k(x)$ are defined on a certain set, for instance, on a closed interval $[a, b]$.

Definition 1₁. *A functional series* (8.8) *is said to be* ***convergent*** *if the sequence of its partial sums*

$$S_n(x) = \sum_{k=1}^{n} u_k(x), \quad n = 1, 2, \ldots \tag{8.9}$$

converges.

The limit

$$S(x) = \lim_{n\to+\infty} S_n(x) \tag{8.10}$$

of the partial sums is called the ***sum of series*** (8.8). *If series* (8.8) *converges and its sum is equal to* $S(x)$ *we write*

$$S(x) = \sum_{k=1}^{\infty} u_k(x) \tag{8.11}$$

Definition 2₁. *A convergent functional series* (8.11) *is said to converge* ***uniformly*** *to its sum* $S(x)$ *on an interval* $[a, b]$ *if the sequence of the corresponding partial sums* $S_n(x)$ *is uniformly convergent to the sum* $S(x)$ *on* $[a, b]$, *i.e. if for every* $\varepsilon > 0$ *there exists* $N = N(\varepsilon)$ *such that the difference between* $S_n(x)$ *and* $S(x)$ *satisfies the inequality*

$$|S(x) - S_n(x)| = \left|\sum_{k=n+1}^{\infty} u_k(x)\right| < \varepsilon \tag{8.12}$$

for all $x \in [a, b]$ *simultaneously when* $n > N(\varepsilon)$ *or, in other words, if*

$$\sup_{a \leqslant x \leqslant b} |S(x) - S_n(x)| = \sup_{a \leqslant x \leqslant b} \Big| \sum_{k=n+1}^{\infty} u_k(x) \Big| \to 0 \quad \textit{as } n \to +\infty \quad (8.12')$$

Examples of uniformly (nonuniformly) convergent functional series can be easily constructed on the basis of uniformly (nonuniformly) convergent sequences. For, if we are given a functional sequence

$$f_1(x),\ f_2(x),\ \ldots,\ f_n(x),\ \ldots \quad (8.13)$$

we can take the series

$$f_1(x) + [f_2(x) - f_1(x)] + [f_3(x) - f_2(x)] + \ldots$$
$$\ldots + [f_n(x) - f_{n-1}(x)] + \ldots \quad (8.14)$$

for which (8.13) is the sequence of its partial sums. Therefore, if sequence (8.13) converges uniformly (nonuniformly) then, according to Definition 2_1, series (8.14) is uniformly (nonuniformly) convergent.

It should be noted that, conversely, uniform (nonuniform) convergence of series (8.14) implies, by virtue of Definition 2_1, uniform (nonuniform) convergence of sequence (8.13).

The following two assertions are an immediate consequence of the definition of uniform convergence:

(1) *The sum of a finite number of uniformly convergent sequences (series) is a uniformly convergent sequence (series).*

(2) *If all the terms of a uniformly convergent sequence (series) are multiplied by a bounded function* $\varphi(x)$ *(in particular, by a constant) this does not affect the character of its convergence which remains uniform.*

The assertions are easily proved. For instance, the second one is proved as follows. Let $f_n(x) \rightrightarrows f(x)$ on $[a, b]$ and let C, $0 < C < +\infty$ be a constant such that $|\varphi(x)| < C$ for all $x \in [a, b]$. Suppose we are given an arbitrary $\varepsilon > 0$. By the uniform convergence of $f_n(x)$ to $f(x)$, we can find $N(\varepsilon)$ such that $|f_n(x) - f(x)| < \frac{\varepsilon}{C}$ for all $x \in [a, b]$ simultaneously when $n > N(\varepsilon)$. But for such values of $n > N(\varepsilon)$ we then have

$$|\varphi(x) f_n(x) - \varphi(x) f(x)| = |\varphi(x)| \cdot |f_n(x) - f(x)| <$$
$$< C \cdot \frac{\varepsilon}{C} = \varepsilon$$

for all $x \in [a, b]$ and hence $\varphi(x) f_n(x) \rightrightarrows \varphi(x) f(x)$, as $n \to \infty$, on the interval $[a, b]$. Regarding $f_n(x)$ as being the nth partial sum of a uniformly convergent functional series and $f(x)$ as the sum of this series we thus conclude that the assertion is valid for the uniformly convergent series as well.

2. Tests for Uniform Convergence. If the limit $f(x)$ of functional sequence (8.13) is known its uniform convergence can be tested on

the basis of Definitions 2 and 2_1 or by means of the corresponding geometric interpretation as it was done when investigating examples 1-4.

But it sometimes turns expedient to reduce the question of uniform convergence of functional sequence (8.13) to testing uniform convergence of corresponding functional series (8.14) for which the sequence in question is the sequence of partial sums. Such a reduction may be useful because there are various tests for uniform convergence of series convenient for practical application.

One of the simplest and most commonly used tests of this kind is the so-called *Weierstrass* M-test* based on comparing a given functional series with a number series having nonnegative terms.

A number series

$$\sum_{k=1}^{+\infty} M_k = M_1 + M_2 + \ldots + M_k + \ldots \tag{8.15}$$

with nonnegative terms is said to be a ***dominant series*** *for a functional series*

$$\sum_{k=1}^{+\infty} u_k(x) = u_1(x) + u_2(x) + \ldots + u_k(x) + \ldots \tag{8.16}$$

on an interval $a \leqslant x \leqslant b$ *if the inequalities*

$$|u_k(x)| \leqslant M_k, \quad k = 1, 2, \ldots \tag{8.17}$$

hold for all $x \in [a, b]$ *simultaneously. Series* (8.16) *is then spoken of as a* ***dominated series.***

Weierstrass' M-test. *If, for functional series* (8.16) *defined on an interval* $[a, b]$, *there exists a convergent dominant series of type* (8.16) *the functional series is uniformly convergent on* $[a, b]$.

Proof. Let an arbitrary $\varepsilon > 0$ be given. Dominant series (8.15) being convergent, we have the inequality

$$\sum_{k=n+1}^{+\infty} M_k < \varepsilon$$

for all sufficiently large n. By relations (8.17), for all such n the inequalities

$$\left|\sum_{k=n+1}^{+\infty} u_k(x)\right| \leqslant \sum_{k=n+1}^{+\infty} |u_k(x)| \leqslant \sum_{k=n+1}^{+\infty} M_k < \varepsilon \tag{8.18}$$

* Weierstrass, Karl Theodor Wilhelm (1815-1897), a prominent German mathematician.

hold for all $x \in [a, b]$ simultaneously. Inequalities (8.18) indicate that series (8.16) is uniformly convergent and thus Weierstrass' test has been proved.

Examples

1. The series $\sum_{n=1}^{+\infty} \frac{\sin nx}{n^2}$ is uniformly convergent on the entire x-axis $-\infty < x < +\infty$ because it can be dominated by the convergent positive series $\sum_{n=1}^{+\infty} \frac{1}{n^2}$ since we have

$$\left|\frac{\sin nx}{n^2}\right| \leqslant \frac{1}{n^2} \text{ for } -\infty < x < +\infty$$

2. Consider the series $\sum_{n=1}^{+\infty} \frac{x}{1+n^4x^2}$ on the positive half of the x-axis $0 \leqslant x < +\infty$. Applying the well known techniques of differential calculus we find $\max\limits_{0\leqslant x<+\infty} \frac{x}{1+n^4x^2} = \frac{1}{2n^2}$. Consequently, $\left|\frac{x}{1+n^4x^2}\right| \leqslant \frac{1}{2n^2}$ for $0 \leqslant x < +\infty$. The series $\sum_{n=1}^{+\infty} \frac{1}{n^2}$ being convergent, we thus conclude, by Weierstrass' test, that the series $\sum_{n=1}^{+\infty} \frac{x}{1+n^4}$ uniformly converges on the positive x-axis $0 \leqslant x < +\infty$.

3. There is no convergent dominant number series for the series $\sum_{n=1}^{+\infty} \frac{(-1)^n}{x+n}$, $0 \leqslant x < +\infty$, since $\max\limits_{0\leqslant x<+\infty} \left|\frac{(-1)^n}{x+n}\right| = \frac{1}{n}$ and the series $\sum_{n=1}^{+\infty} \frac{1}{n}$ (the so-called *harmonic series*) is divergent. But, by Leibniz'* test (e.g. see [8], Chapter 13, § 5, inequality (13.80)), the inequality

$$\left|\sum_{k=n}^{+\infty} \frac{(-1)^k}{x+k}\right| \leqslant \frac{1}{x+n} \leqslant \frac{1}{n}$$

holds for every $x \in [0, +\infty]$ and consequently the definition of uniform convergence of a functional series (see relation (8.12))

* Leibniz, Gottfried Wilhelm (1646-1716), the great German philosopher and mathematician.

implies that the series $\sum_{n=1}^{+\infty} \frac{(-1)^n}{x+n}$ converges uniformly on the entire positive half of the x-axis $0 \leqslant x < +\infty$. This example shows that Weierstrass' test provides only a sufficient condition for a functional series to be uniformly convergent but not a necessary one.

Now let us proceed to formulate the basic (*Cauchy*) *criterion for uniform convergence* which plays an important theoretical role because, unlike Weierstrass' test, it gives us a necessary and sufficient condition for uniform convergence and enables us to establish more subtle sufficient conditions (compared with Weierstrass' test).

Cauchy's* Test (for Uniform Convergence of a Sequence). *For a functional sequence $\{f_n(x)\}$ to be uniformly convergent to a function $f(x)$ on an interval $[a, b]$ it is necessary and sufficient that for every $\varepsilon > 0$ there exist $N = N(\varepsilon)$ such that for each $n > N(\varepsilon)$ and all $p > 0$ the inequality*

$$|f_{n+p}(x) - f_n(x)| < \varepsilon \tag{8.19}$$

should hold for all $x \in [a, b]$ simultaneously.

Proof. Necessity. Let $f_n(x) \rightrightarrows f(x)$ on $[a, b]$. Then, given an arbitrary $\varepsilon > 0$, there is $N(\varepsilon)$ such that for all $n > N(\varepsilon)$ and all $p > 0$ the inequalities

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2} \quad \text{and} \quad |f_{n+p}(x) - f(x)| < \frac{\varepsilon}{2}$$

hold for all $x \in [a, b]$. Therefore we have $|f_{n+p}(x) - f_n(x)| \leqslant$ $\leqslant |f_{n+p}(x) - f(x)| + |f(x) - f_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ for all $n > N(\varepsilon)$, all $p > 0$ and all $x \in [a, b]$.

Sufficiency. If inequality (8.19) is fulfilled for all $x \in [a, b]$ it follows that for every fixed $x \in [a, b]$ the numerical sequence $f_n(x)$, $n = 1, 2, \ldots$, is convergent because it is a Cauchy (fundamental) sequence. Hence, the functional sequence $f_n(x)$, $n = 1, 2, \ldots$, converges on the entire interval $[a, b]$. Let the limit function be denoted by $f(x)$. Passing to the limit in inequality (8.19) as $p \to +\infty$ we find that

$$|f_n(x) - f(x)| \leqslant \varepsilon$$

for all $n > N(\varepsilon)$ and for all $x \in [a, b]$ simultaneously. But this implies that $f_n(x) \rightrightarrows f(x)$ on the interval $[a, b]$ and Cauchy's test has thus been proved.

* Cauchy, Augustin Louis (1789-1857), a famous French mathematician.

Applying Cauchy's test for uniform convergence of a sequence to the sequence of partial sums

$$S_1(x) = u_1(x), \quad S_2(x) = u_1(x) + u_2(x), \ \ldots, \ S_n(x) = \\ = u_1(x) + \ldots + u_n(x), \ \ldots$$

of a functional series $\sum\limits_{k=1}^{+\infty} u_k(x)$ we arrive at

Cauchy's Test (for Uniform Convergence of a Series). *A functional series*

$$\sum_{k=1}^{+\infty} u_k(x) = u_1(x) + u_2(x) + \ldots + u_k(x) + \ldots \tag{8.20}$$

is uniformly convergent on an interval $[a, b]$ *if and only if for every* $\varepsilon > 0$ *there exists* $N = N(\varepsilon)$ *such that for each* $n > N(\varepsilon)$ *and all* $p > 0$ *the inequality*

$$|S_{n+p}(x) - S_n(x)| = \left|\sum_{k=n+1}^{n+p} u_k(x)\right| = |u_{n+1}(x) + \ldots + u_{n+p}(x)| < \varepsilon \tag{8.21}$$

is simultaneously fulfilled for all $x \in [a, b]$.

On the basis of Cauchy's test we can establish the following

Abel's* Test (for Uniform Convergence of a Series). *If the partial sums of a series*

$$\sum_{k=1}^{+\infty} u_k(x) = u_1(x) + u_2(x) + \ldots + u_k(x) + \ldots \tag{8.22}$$

are uniformly bounded on an interval $[a, b]$, *i.e. if there is a constant* C, $0 < C < +\infty$, *such that*

$$|S_n(x)| = \left|\sum_{k=1}^{n} u_k(x)\right| < C \quad \textit{for } n = 1, 2, \ldots \tag{8.23}$$

for all $x \in [a, b]$, *and if*

$$\alpha_1(x), \ \alpha_2(x), \ \ldots \ \alpha_k(x), \ \ldots \tag{8.24}$$

is a monotone nonincreasing functional sequence uniformly convergent to zero on the interval $[a, b]$, *the series*

$$\sum_{k=1}^{+\infty} \alpha_k(x)\, u_k(x) \tag{8.25}$$

uniformly converges on $[a, b]$.

---

* Abel, Niels Henrik (1802-1829), a famous Norwegian mathematician.

Before proving Abel's test let us consider an example of its application.

4. We take the series $\sum_{k=1}^{+\infty} \frac{\sin kx}{k}$ which can be regarded as the result of multiplying the terms of the series

$$\sum_{k=1}^{+\infty} \sin kx = \sin x + \sin 2x + \ldots + \sin kx + \ldots \tag{8.26}$$

by the members of the sequence

$$1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{k}, \ldots \tag{8.27}$$

We have the inequality

$$|S_n(x)| \leqslant \frac{1}{\left|\sin \frac{x}{2}\right|} \quad (x \neq 2\pi m, \ m = 1, 2, \ldots)$$

for the partial sums of series (8.26) (e.g. see [8], Chapter 13, § 5) and consequently the inequalities

$$|S_n(x)| \leqslant \frac{1}{\sin \frac{\alpha}{2}} = \text{const} < +\infty, \ n = 1, 2, \ldots$$

simultaneously hold for all x satisfying the conditions

$$2m\pi + \alpha \leqslant x \leqslant (2m+1)\pi - \alpha, \quad 0 < \alpha < \pi,$$
$$m = 0, \pm 1, \pm 2, \ldots \tag{8.28}$$

Since (8.27) is a monotone decreasing number sequence converging to zero it can be thought of as a uniformly convergent functional sequence satisfying the conditions enumerated in Abel's theorem. Thus, on every interval determined by conditions (8.28) the series $\sum_{k=1}^{+\infty} \frac{\sin kx}{k}$ satisfies the requirements of Abel's test and hence the series is uniformly convergent on the interval.

Proof of Abel's test. To accomplish the proof we shall show that under the above assumptions series (8.25) satisfies the condition of Cauchy's test for uniform convergence. We have

$$\begin{aligned} &\alpha_{n+1}u_{n+1} + \alpha_{n+2}u_{n+2} + \ldots + \alpha_{n+p}u_{n+p} = \\ &= \alpha_{n+1}[S_{n+1} - S_n] + \alpha_{n+2}[S_{n+2} - S_{n+1}] + \\ &\ldots + \alpha_{n+p}[S_{n+p} - S_{n+p-1}] = -\alpha_{n+1}S_n + \\ &+ (\alpha_{n+1} - \alpha_{n+2}) S_{n+1} + \ldots \\ &\ldots + (\alpha_{n+p-1} - \alpha_{n+p}) S_{n+p-1} + \alpha_{n+p}S_{n+p} \end{aligned} \tag{8.29}$$

Making use of the inequalities $\alpha_1(x) \geqslant \alpha_2(x) \geqslant \ldots \geqslant \alpha_n(x) \geqslant \geqslant \alpha_{n+1}(x) \geqslant \ldots$ and the relation $|S_n(x)| \leqslant C$ (fulfilled for all $n = 1, 2, \ldots$ and all $x \in [a, b]$) we deduce from equality (8.29) the relation

$$|\alpha_{n+1}u_{n+1} + \ldots + \alpha_{n+p}u_{n+p}| \leqslant$$
$$\leqslant C\{\alpha_{n+1} + (\alpha_{n+1} - \alpha_{n+2}) + (\alpha_{n+2} - \alpha_{n+3}) + \ldots$$
$$\ldots + (\alpha_{n+p-1} - \alpha_{n+p}) + \alpha_{n+p}\} = 2C\alpha_{n+1} \leqslant 2C\varepsilon_{n+1}$$

where $\varepsilon_{n+1} = \sup\limits_{a \leqslant x \leqslant b} \alpha_{n+1}(x) \to 0$ as $n \to \infty$, which holds for all $p > 0$ and for all $x \in [a, b]$ simultaneously. Hence, under the assumptions given in Abel's test, series (8.25) satisfies the condition of Cauchy's test.

§ 2. PROPERTIES OF UNIFORMLY CONVERGENT FUNCTIONAL SEQUENCES AND SERIES

1. Continuity and Uniform Convergence.

Theorem 8.1. (A) *If a sequence of continuous functions $f_1(x)$, $f_2(x), \ldots, f_n(x), \ldots$ uniformly converges on an interval $[a, b]$ to a function $f(x)$ the limit function is also continuous on $[a, b]$.*

(B) *If all the terms of a series*

$$S(x) = \sum_{k=1}^{+\infty} u_k(x) \tag{8.30}$$

are continuous functions on an interval $[a, b]$ and the series is uniformly convergent on $[a, b]$ its sum $S(x)$ is also continuous on the interval.

Proof. (A) Take an arbitrary point $x \in [a, b]$ and let $(x + h) \in \in [a, b]$. We shall establish the continuity of $f(x)$ at the point x. For this purpose we estimate the difference $f(x + h) - f(x)$. Suppose we are given an arbitrary $\varepsilon > 0$. Let us show that for all the values of h sufficiently small in their moduli the modulus of the difference is smaller than ε. We have

$$|f(x+h) - f(x)| \leqslant |f(x+h) - f_n(x+h)| +$$
$$+ |f_n(x+h) - f_n(x)| + |f_n(x) - f(x)| \tag{8.31}$$

Taking a sufficiently large n we obtain, by the uniform convergence of $f_n(x)$ to $f(x)$ on $[a, b]$, the inequalities

$$|f(x+h) - f_n(x+h)| < \frac{\varepsilon}{3} \quad \text{for all } (x+h) \in [a, b] \tag{8.32}$$

and

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3} \quad \text{for all } x \in [a, b] \tag{8.33}$$

We now fix the above-chosen value of n and consider the term $|f_n(x+h) - f_n(x)|$ entering into the right-hand side of inequality

(8.31). Since $f_n(x)$ is a continuous function, there is $\delta = \delta(\varepsilon) > 0$ such that for all h satisfying the inequality $|h| < \delta(\varepsilon)$ we have

$$|f_n(x+h) - f_n(x)| < \frac{\varepsilon}{3} \tag{8.34}$$

But then, by virtue of (8.32), (8.33) and (8.34), we derive from (8.31) the relation $|f(x+h) - f(x)| < \varepsilon$ valid for all h satisfying the condition $|h| < \delta(\varepsilon)$, which means that $f(x)$ is continuous at the point x arbitrarily chosen on the interval $[a, b]$. Hence, $f(x)$ is continuous at each point $x \in [a, b]$, that is the function $f(x)$ is continuous on $[a, b]$.

It should be noted that if x is an end point of the interval $[a, b]$ then x can be given only a nonnegative increment if $x = a$ and only a nonpositive one if $x = b$, and thus the above argument indicates that $f(x)$ is continuous on the left at $x = b$ and on the right at $x = a$.

(B) Every partial sum $S_n(x) = \sum\limits_{k=1}^{n} u_k(x)$ of the series $\sum\limits_{k=1}^{+\infty} u_k(x)$ being a continuous function as a sum of a finite number of continuous functions for any $n = 1, 2, 3, \ldots$, we see that $S(x)$ is continuous because, by the hypothesis, the series is uniformly convergent and hence $S_n(x) \rightrightarrows S(x)$ on $[a, b]$ which implies, by (A), the continuity of $S(x)$. The theorem has thus been proved.

The condition of uniform convergence is only sufficient but not necessary for the limit of a sequence of continuous functions to be continuous. To illustrate this we can take example 4 in Sec. 1 of § 1 in which we considered the sequence $f_n(x) = \frac{2nx}{1+n^2x^2}$, $n = 1, 2, \ldots$, whose members are continuous functions. As was shown, the sequence converges nonuniformly to the continuous function $f(x) \equiv 0$ on the positive part of the x-axis $0 \leqslant x < +\infty$.

But there is a special class of sequences and series for which uniform convergence, as was proved by Dini*, is equivalent to continuity of the limit of the sequence or of the sum of the series.

Theorem 8.1' (Dini's Theorem on Uniform Convergence). (A) *If a sequence of continuous functions $f_n(x)$, $n = 1, 2, \ldots$, defined on $[a, b]$** is nondecreasing, that is $f_1(x) \leqslant f_2(x) \leqslant \ldots \leqslant f_n(x) \leqslant \ldots$ on $[a, b]$, and if $f_n(x)$ converges to a continuous function $f(x)$ the convergence is uniform on $[a, b]$.*

---

* Dini, Ulisse (1845-1918), an Italian mathematician.

** The condition that $[a, b]$ is a bounded and closed interval is essentially used in the proof of this theorem which also remains true if $[a, b]$ is replaced by an arbitrary bounded closed set X.

(B) *If the sum of a series* $S(x) = \sum_{k=1}^{+\infty} u_k(x)$ *with nonnegative continuous terms defined in an interval* $[a, b]$ *is continuous on* $[a, b]$ *the series is uniformly convergent on the interval.*

Proof. (A) Let us show that for any $\varepsilon > 0$ there is n such that

$$0 \leqslant R_n(x) = f(x) - f_n(x) < \varepsilon \tag{8.35}$$

for all $x \in [a, b]$ simultaneously. Then, the sequence $R_n(x)$, $n = 1, 2, \ldots$, being obviously monotone, that is

$$R_1(x) \geqslant R_2(x) \geqslant \ldots \geqslant R_n(x) \geqslant \ldots \tag{8.36}$$

relation (8.35) must also hold for all sufficiently large n, which implies the uniform convergence.

We shall prove the above assertion by contradiction. Suppose that for a certain $\varepsilon_0 > 0$ there is no such n. Then for each $n = 1, 2, \ldots$ there is $x_n \in [a, b]$ such that

$$R_n(x_n) \geqslant \varepsilon_0 \tag{8.37}$$

Applying the Bolzano*-Weierstrass theorem to the sequence of the points $x_1, x_2, \ldots, x_n, \ldots$ belonging to the interval $[a, b]$ we can assert that there is a subsequence $x_{n_1}, x_{n_2}, \ldots, x_{n_k}, \ldots$ convergent to a point $x_0 \in [a, b]$. The function $R_n(x) = f(x) - f_n(x)$ (the difference of two continuous functions) is continuous and therefore we can write, for every fixed m, the relation

$$\lim_{n_k \to +\infty} R_m(x_{n_k}) = R_m(x_0)$$

But for every m and any sufficiently large k we have $n_k > m$ and consequently, by virtue of (8.36) and (8.37), we obtain

$$R_m(x_{n_k}) \geqslant R_{n_k}(x_{n_k}) \geqslant \varepsilon_0$$

Passing to the limit in the last inequality as $n_k \to +\infty$ we see that $R_m(x_0) \geqslant \varepsilon_0$ for any m. But this contradicts the relation $\lim_{m \to \infty} R_m(x_0) = 0$ implied by the convergence of $f_m(x)$ to $f(x)$ at the point x_0.

(B) The partial sums $S_n(x) = \sum_{k=1}^{n} u_k(x)$, $n = 1, 2, \ldots$, with nonnegative continuous terms $u_k(x)$ form a nondecreasing sequence of continuous functions convergent, by the hypothesis, to a continuous function $S(x)$. Therefore, by (A), the sequence converges uniformly and thus the series is also uniformly convergent.

*Bolzano, Bernard (1781-1848), an Italian mathematician.

2. Passage to Limit Under the Sign of Integration and Termwise Integration of a Series. If

$$\lim_{n\to+\infty}\int_{x_0}^{x} f_n(\xi)\,d\xi=\int_{x_0}^{x}[\lim_{n\to+\infty} f_n(\xi)]\,d\xi \tag{8.38$_1$}$$

or

$$\int_{x_0}^{x}\left\{\sum_{k=1}^{+\infty} u_k(\xi)\right\}d\xi=\sum_{k=1}^{+\infty}\int_{x_0}^{x} u_k(\xi)\,d\xi \tag{8.38$_2$}$$

we say, accordingly, that it is permissible *to pass to the limit under the integral sign* in the integral $\int_{x_0}^{x} f_n(\xi)\,d\xi$ or that the series $\sum_{k=1}^{+\infty} u_k(x)$ admits *termwise (term-by-term) integration from* x_0 to x.

Relation (8.38$_2$) can be regarded as a generalization of the theorem on an integral of a sum to the case of an infinite number of summands.

Replacing the functional sequence $\{f_n(x)\}$ by a series $\sum_{k=1}^{+\infty} u_k(x)$ for which it serves as the sequence of partial sums or, conversely, replacing the series $\sum_{k=1}^{+\infty} u_k(x)$ by the sequence of its partial sums we can easily transform relation (8.38$_1$) to the form (8.32$_2$) and (8.38$_2$) to (8.38$_1$).

Hence. when investigating conditions for relation (8.38$_1$) to be valid we incidentally obtain the answer to the question of validity of relation (8.38$_2$) and vice versa. It should be noted that, for relations (8.38$_1$) and (8.38$_2$) to be true, the existence of the integrals and convergence of the corresponding sequences and series are not sufficient. This can be confirmed by the examples considered at the end of the present section. For these relations to hold, an additional condition should be imposed. It turns out that such a sufficient condition is (1) *uniform convergence* (this will be proved below) or (2) *convergence in the mean* (which will be shown in § 6 of the present chapter).

Theorem 8.2. (A) *If a sequence of continuous functions* $\{f_n(x)\}$ *uniformly converges on an interval* $[a, b]$ *to a function* $f(x)$, *i.e.*

$$f_n(x)\rightrightarrows f(x)\quad on\quad [a, b] \tag{8.39}$$

the sequence of integrals $\{\int_{x_0}^{x} f_n(z)\,dz\}$ *uniformly converges (as a sequence*

of functions dependent on x) *on the interval* $[a, b]$ *to the integral* $\int_{x_0}^{x} f(z)\,dz$, *i.e.*

$$\int_{x_0}^{x} f_n(z)\,dz \rightrightarrows \int_{x_0}^{x} f(z)\,dz \tag{8.40}$$

for any $x_0 \in [a, b]$.

(B) *If a series*

$$S(x) = \sum_{k=1}^{+\infty} u_k(x) \tag{8.41}$$

whose terms are continuous on an interval $[a, b]$ *is uniformly convergent on the interval we have the relation*

$$\int_{x_0}^{x} S(z)\,dz = \sum_{k=1}^{+\infty} \int_{x_0}^{x} u_k(z)\,dz \tag{8.42}$$

which means that series (8.41) *can be integrated term-by-term within the limits from* x_0 *to* x *for any* x_0 *and* x *belonging to the interval* $[a, b]$ *and the series on the right-hand side of* (8.42) *uniformly converges* (*with respect to* x) *on the interval* $[a, b]$ *for any* $x_0 \in [a, b]$.

Proof. (A) Let there be given an arbitrary $\varepsilon > 0$. Take $N(\varepsilon)$ such that the inequality

$$|f_n(x) - f(x)| < \frac{\varepsilon}{b-a} \tag{8.43}$$

holds for all $x \in [a, b]$ simultaneously when $n > N(\varepsilon)$, the choice of $N(\varepsilon)$ being possible by relation (8.34). According to Theorem 8.1, the function $f(x)$ is continuous as a limit of a uniformly convergent sequence of continuous functions. Therefore the integral $\int_{x_0}^{x} f(z)\,dz$ exists for any x_0 and x belonging to $[a, b]$. Let us estimate the difference $\int_{x_0}^{x} f_n(z)\,dz - \int_{x_0}^{x} f(z)\,dz$. By (8.43), we have, for every $n > N(\varepsilon)$, the inequality

$$\left| \int_{x_0}^{x} f_n(z)\,dz - \int_{x_0}^{x} f(z)\,dz \right| = \left| \int_{x_0}^{x} [f_n(z) - f(z)]\,dz \right| \leqslant$$

$$\leqslant \int_{x_0}^{x} |f_n(z) - f(z)|\,dz \leqslant |x - x_0| \frac{\varepsilon}{b-a} \leqslant \varepsilon \tag{8.44}$$

which implies (8.40). It follows from relation (8.40) that

$$\lim_{n\to+\infty}\int_{x_0}^{x} f_n(z)\,dz=\int_{x_0}^{x} f(z)\,dz \quad \text{for } x,\ x_0\in[a,\ b] \tag{8.45}$$

Thus, if a sequence of continuous functions $\{f_n(x)\}$ is uniformly convergent on $[a,\ b]$ it is allowable to take the limit behind the integral sign in the integral $\int_{x_0}^{x} f_n(z)\,dz$ for any x_0 and x belonging to $[a,\ b]$.

(B) The partial sum $S_n(x)=\sum_{k=1}^{n} u_k(x)$ is continuous for every $n=1,\ 2,\ \ldots$ as a sum of a finite number of continuous functions. By the hypothesis, we have

$$S_n(x) \rightrightarrows S(x) \quad \text{on} \quad [a,\ b]$$

Then, by (A), we conclude that

$$\int_{x_0}^{x} S_n(z)\,dz \rightrightarrows \int_{x_0}^{x} S(z)\,dz \quad \text{on } [a,\ b]$$

Observe that

$$\int_{x_0}^{x} S_n(z)\,dz=\int_{x_0}^{x}\sum_{k=1}^{n} u_k(z)\,dz=\sum_{k=1}^{n}\int_{x_0}^{x} u_k(z)\,dz \tag{8.46}$$

and therefore relation (8.45) can be rewritten as

$$\left\{\sum_{k=1}^{n}\int_{x_0}^{x} u_k(z)\,dz\right\} \rightrightarrows \int_{x_0}^{x} S(z)\,dz \quad \text{on } [a,\ b] \tag{8.47}$$

where the expression in the curly brackets is the nth partial sum of series (8.42). Consequently, equality (8.42) is true and the series on its right-hand side is uniformly convergent on $[a,\ b]$, which is what we set out to prove.

Note. The theorem also remains true when the functions $f_n(x)$, $n=1,\ 2,\ \ldots$, may have discontinuities but are integrable. In such a case the function $f(x)$ is also integrable and relation (8.40) is fulfilled provided that $\{f_n(x)\}$ converges uniformly.

The condition of uniform convergence is only sufficient but not necessary for a series to admit term-by-term integration and for the passage to the limit under the sign of integration to be permissible. For instance, the sequence $f_n(x)=x^n$ converges nonuniformly on the interval $0\leqslant x\leqslant 1$ to its limit

$$f(x)=\begin{cases} 0 \text{ for } 0\leqslant x\leqslant 1 \\ 1 \text{ for } x=1 \end{cases}$$

But at the same time

$$\int_{x_0}^{x} f_n(z)dz = \int_{x_0}^{x} z^n\,dz = \frac{x^{n+1} - x_0^{n+1}}{n+1} \to 0 = \int_{x_0}^{x} f(z)\,dz \quad \text{for } n \to +\infty$$

and any x_0 and x belonging to the closed interval $[0, 1]$. On the other hand, there are cases when a nonuniformly convergent sequence of integrable functions $f_n(x)$ converges to its limit in such a way that $\lim\limits_{n\to+\infty} \int_{x_0}^{x} f_n(z)\,dz \neq \int_{x_0}^{x} f(z)\,dz$. For example, we have, for $n \to +\infty$, the relation

$$f_n(x) = 4nx^3e^{-nx^4} \to f(x) \equiv 0, \quad -\infty < x < +\infty$$

but

$$\int_0^1 4nx^3e^{-nx^4}\,dx = 1 - e^{-n} \nrightarrow \int_0^1 0\cdot dx = 0 \quad \text{as } n \to +\infty$$

3. Passage to Limit Under the Sign of Differentiation and Termwise Differentiation of a Series. If

$$\lim_{n\to+\infty} f_n'(x) = \{\lim_{n\to+\infty} f_n(x)\}' \tag{8.48}$$

or

$$\left\{\sum_{k=1}^{+\infty} u_k(x)\right\}' = \sum_{k=1}^{+\infty} u_k'(x) \tag{8.49}$$

we say, accordingly, that it is permissible *to pass to the limit under the differentiation sign* or that the series $\sum\limits_{k=1}^{\infty} u_k(x)$ can be *differentiated term-by-term*.

Relations (8.48) and (8.49) are equivalent in the same sense as relations (8.38_1) and (8.38_2).

Formula (8.49) can be regarded as a generalization of the rule for differentiating a sum to the case of an infinite number of summands.

It turns out that the condition of existence of the derivatives and convergence of the corresponding sequences and series is not sufficient for relations (8.48) and (8.49) to be valid and that some additional requirements should be imposed to guarantee the validity. Sufficient conditions of this kind are given by

Theorem 8.3. (A) *If a sequence of continuously differentiable functions** $\{f_n(x)\}$ *converges to* $f(x)$ *on* $[a, b]$, *i.e.*

$$f_n(x) \to f(x), \quad x \in [a, b] \tag{8.50}$$

---

* A function $f(x)$ is said to be continuously differentiable if it is differentiable and its derivative is a continuous function.

and the sequence $\{f'_n(x)\}$ *of their derivatives is uniformly convergent to a function* $\varphi(x)$ *on* $[a, b]$, *i.e.*

$$f'_1(x),\ f'_2(x),\ \ldots,\ f'_n(x),\ \ldots\ \rightrightarrows\ \varphi(x)\quad \text{on}\quad [a, b] \tag{8.51}$$

the function $f(x)$ *is also differentiable on* $[a, b]$ *and*

$$f'(x) = \varphi(x) = \lim_{n\to+\infty} f'_n(x) \tag{8.52}$$

Hence, under these conditions, it is permissible to pass to the limit under the differentiation sign.

(B) *If a series*

$$S(x) = \sum_{k=1}^{+\infty} u_k(x) \tag{8.53}$$

with continuously differentiable terms converges on $[a, b]$ *and the series*

$$\sigma(x) = \sum_{k=1}^{+\infty} u'_k(x) \tag{8.54}$$

whose terms are the derivatives of $u_k(x)$, $k = 1, 2, \ldots$, *converges uniformly on* $[a, b]$ *the sum* $S(x)$ *of the former series is differentiable on the interval* $[a, b]$ *and the equality*

$$S'(x) = \sigma(x) = \sum_{k=1}^{+\infty} u'_k(x) \tag{8.55}$$

is fulfilled at each point of the interval. Thus, under the above assumptions, series (8.53) *can be differentiated termwise.*

Proof. (A) By the hypothesis, the derivatives $f'_n(x)$ being continuous functions and the convergence being uniform on $[a, b]$ ($f'_n(x) \rightrightarrows \varphi(x)$), we conclude, on the basis of Theorem 8.2, that

$$\lim_{n\to+\infty} \int_{x_0}^{x} f'_n(z)\,dz = \int_{x_0}^{x} \varphi(z)\,dz \tag{8.56}$$

i.e.

$$\lim_{n\to+\infty} [f_n(x) - f_n(x_0)] = \int_{x_0}^{x} \varphi(z)\,dz \tag{8.57}$$

Passing to the limit in the left-hand side of equality (8.57) we obtain

$$f(x) - f(x_0) = \int_{x_0}^{x} \varphi(z)\,dz$$

and consequently

$$f(x) = f(x_0) + \int_{x_0}^{x} \varphi(z)\,dz \tag{8.58}$$

Hence we see that the function $f(x)$ is differentiable because it is equal to the sum of the constant $f(x_0)$ and the integral $\int\limits_{x_0}^{x} \varphi(z)\,dz$ which are differentiable functions. Now differentiating both sides of equality (8.58) with respect to x we receive

$$f'(x) = \varphi(x) = \lim_{n\to+\infty} f'_n(x)$$

(B) Putting $S_n(x) = \sum\limits_{k=1}^{n} u_k(x)$ we can write, by the hypothesis,

$$S_n(x) \to S(x) \quad \text{on} \quad [a, b]$$

and

$$S'_n(x) \rightrightarrows \sigma(x) \quad \text{on } [a, b]$$

and the function $S'_n(x) = \sum\limits_{k=1}^{n} u'_k(x)$ is continuous on $[a, b]$ for every $n = 1, 2, \ldots$. Consequently, by (A), $S(x)$ is a differentiable function on the interval $[a, b]$ and the relation

$$S'(x) = \sigma(x) = \sum_{k=1}^{+\infty} u'_k(x)$$

holds everywhere on $[a, b]$, which is what we set out to prove.

If a sequence of derivatives converges nonuniformly equality (8.52) is not necessarily fulfilled. For instance, we have

$$f_n(x) = \frac{1}{n}\ln\left(nx + \sqrt{n^2x^2+1}\right) \to$$

$$\to f(x) \equiv 0 \quad \text{for } n\to+\infty, \quad -\infty < x < +\infty$$

but at the same time

$$\lim_{n\to+\infty} f'_n(0) = \lim_{n\to+\infty}\left(\frac{1}{\sqrt{n^2x^2+1}}\bigg|_{x=0}\right) = 1 \neq f'(0) = 0$$

4. Term-by-Term Passage to Limit in Functional Sequences and Series. Generally speaking, the well known theorem on a limit of a sum is not true if the number of summands is infinite. Thus, for instance, every term entering into the series on the right-hand side of the equality

$$x = \frac{2l}{\pi}\sum_{k=1}^{+\infty}\frac{(-1)^{k+1}}{k}\sin\frac{k\pi x}{l}, \quad -l < x < l, \quad l \neq 0$$

(proved in § 2, Sec. 5 of Chapter 11) and the sum of this series tend to a finite limit as $x \to l - 0$. But if we formally apply the theorem on a limit of a sum to the above series for $x \to l - 0$ we arrive at an absurd equality $l = 0$.

But the theorem on a limit of a sum can be extended to the case of an infinite number of summands if some additional restrictions are imposed. Namely, we have

Theorem 8.4. (A) *Let a functional series*

$$S(x)=\sum_{k=1}^{+\infty} u_k(x)=u_1(x)+u_2(x)+\ldots+u_k(x)+\ldots \tag{8.59}$$

converge uniformly in a neighbourhood of a point x_0 and let

$$\lim_{x\to x_0} u_k(x)=c_k \quad \textit{for } k=1,\ 2,\ \ldots \tag{8.60}$$

Then the number series $\sum\limits_{k=1}^{\infty} c_k$ *is convergent and*

$$\lim_{x\to x_0}\sum_{k=1}^{+\infty} u_k(x)=\sum_{k=1}^{+\infty} c_k \tag{8.61}$$

which means that in a uniformly convergent series it is permissible to perform a term-by-term passage to the limit.

(B) *If a functional sequence $f_1(x)$, $f_2(x)$, ..., $f_n(x)$, ... is uniformly convergent in a neighbourhood of a point x_0 and for every n there exists a finite limit*

$$\lim_{x\to x_0} f_n(x)=A_n$$

the numerical sequence A_1, A_2, ..., A_n, ... is also convergent and

$$\lim_{x\to x_0}\ \lim_{n\to+\infty} f_n(x)=\lim_{n\to+\infty}\ \lim_{x\to x_0} f_n(x)$$

Proof. Let an arbitrary $\varepsilon>0$ be given. Series (8.59) being uniformly convergent in a neighbourhood of x_0, there is $N(\varepsilon)$ such that for all $n>N(\varepsilon)$ and all $p>0$ the inequality

$$|u_{n+1}(x)+\ldots+u_{n+p}(x)|<\varepsilon \tag{8.62}$$

holds for all x belonging to the neighbourhood. Passing to the limit in inequality (8.62), as $x\to x_0$, we obtain the inequality

$$|c_{n+1}+\ldots+c_{n+p}|\leqslant\varepsilon \tag{8.63}$$

which is valid for all $n>N(\varepsilon)$ and all $p>0$. Consequently the series $\sum\limits_{k=1}^{\infty} c_k$ is convergent. Now, making the index p in (8.62) and (8.63) tend to infinity we derive the inequalities

$$\Big|\sum_{k=n+1}^{\infty} c_k\Big|\leqslant\varepsilon, \quad \Big|\sum_{k=n+1}^{\infty} u_k(x)\Big|\leqslant\varepsilon \tag{8.64}$$

which are fulfilled for all $n > N(\varepsilon)$ and for all x belonging to the chosen neighbourhood of x_0. Let us fix an arbitrary $n > N(\varepsilon)$ and take $\delta = \delta(\varepsilon)$ such that the condition

$$\left|\sum_{k=1}^{n} u_k(x) - \sum_{k=1}^{n} c_k\right| < \varepsilon \quad \text{for } 0 < |x - x_0| < \delta(\varepsilon) \tag{8.65}$$

holds. Then, for $0 < |x - x_0| < \delta(\varepsilon)$, we have, by virtue of (8.64) and (8.65), the relation

$$\left|\sum_{k=1}^{+\infty} u_k(x) - \sum_{k=1}^{+\infty} c_k\right| \leqslant \left|\sum_{k=1}^{n} u_k(x) - \sum_{k=1}^{n} c_k\right| +$$
$$+ \left|\sum_{k=n+1}^{+\infty} u_k(x)\right| + \left|\sum_{k=n+1}^{+\infty} c_k\right| \leqslant \varepsilon + \varepsilon + \varepsilon = 3\varepsilon \tag{8.66}$$

and hence part (A) of the theorem has been proved.

(B) This part of the theorem follows from (A) if we take the series

$$f_1(x) + [f_2(x) - f_1(x)] + \ldots + [f_n(x) - f_{n-1}(x)] + \ldots$$

whose partial sums form the sequence $f_1(x), f_2(x), \ldots, f_n(x), \ldots$, all the conditions of part (A) being fulfilled for this series.

§ 3. POWER SERIES

A functional series of the form

$$\sum_{k=0}^{+\infty} c_k x^k = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n + \ldots \tag{8.67}$$

or of the form

$$\sum_{k=0}^{+\infty} c_k (x - x_0)^k = c_0 + c_1 (x - x_0) + c_2 (x - x_0)^2 + \ldots$$
$$\ldots + c_n (x - x_0)^n + \ldots \tag{8.68}$$

where the coefficients $c_0, c_1, \ldots, c_n, \ldots$ are constant numbers is called a power series. A simple change of variable of the form $x' = x - x_0$ reduces series (8.68) to (8.67). Therefore in what follows we shall restrict ourselves to the series of form (8.67). The method of representing a function in the form of a power series or, in other words, *expanding a function into a power series*, is widely applied both in theoretical studies and in approximate calculations. These applications will be discussed in more detail in § 5. Here we are going to investigate the basic properties of power series.

1. Interval of Convergence of Power Series. Radius of Convergence. We shall first investigate the structure of the domain of convergence

of a power series. In contrast to general series whose domain of convergence can be arbitrary sets of any complex structure, the domain of convergence of a power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ is always an interval of the x-axis which can be a closed interval, a half-closed (half-opened) interval or an open interval. This interval may also degenerate into a single point ($x = 0$) or coincide with the whole x-axis. Every power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ converges at the point $x = 0$ since at this point it turns into a numerical series of the form

$$c_0 + c_1 \cdot 0 + c_2 \cdot 0 + \ldots + c_n \cdot 0 + \ldots = c_0$$

There are power series which converge only at the point $x = 0$. For instance, the series $\sum\limits_{n=0}^{+\infty} n! x^n$* is of this kind. Actually, for any $x \neq 0$ we have

$$\lim_{n \to +\infty} \frac{(n+1)! \, |x|^{n+1}}{n! \, |x|^n} = \lim_{n \to +\infty} (n+1) |x| = +\infty$$

and consequently, by the well known D'Alembert's** test, the series $\sum\limits_{n=0}^{+\infty} n! x^n$ is divergent for $x \neq 0$. There are also power series convergent on the whole x-axis, for example, the series $\sum\limits_{n=0}^{+\infty} \frac{x^n}{n!}$ whose convergence can be easily established for any x, with the help of D'Alembert's test. Now let us take an example of a power series with a domain of convergence which neither coincides with the entire x-axis nor degenerates into the point $x = 0$. Such is the series $1 + x + x^2 + \ldots + x^n + \ldots$ (a geometric series) whose terms form a geometric progression with common ratio x. As is known, this series converges for $|x| < 1$ and diverges for $|x| \geqslant 1$. Thus, its domain of convergence is the finite (open) interval $-1 < x < 1$ with centre at the point $x = 0$. The following theorem indicates that the above result is a special case of a general property of power series.

Theorem 8.5. If the domain of convergence of a power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ *neither degenerates into the point* $x = 0$ *nor coincides with*

---

* We remind the reader that, by definition, $0! = 1$.

** D'Alembert, Jean le Rond (1717-1783), a French philosopher and mathematician.

the entire x-axis there is a finite open interval $(-R, R)$, $0 < R < +\infty$ (termed the **interval of convergence of the power series***) such that the series is absolutely convergent at each interior point of the interval and divergent at every point lying outside the closed interval $[-R, R]$.**

The proof of the theorem is based on the following

Lemma. *If a power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ converges for $x = \alpha \neq 0$ it converges absolutely for every x satisfying the condition $|x| < |\alpha|$.*

Proof of the lemma. The series $\sum\limits_{k=0}^{+\infty} c_k \alpha^k$ being convergent, it follows that $c_k \alpha^k \to 0$ as $k \to +\infty$. Therefore there exists $A = \text{const} < +\infty$ such that $|c_k \alpha^k| \leqslant A$ for all $k = 0, 1, 2, \ldots$. Now let $|x| < |\alpha|$. If we put $q = \frac{|x|}{|\alpha|}$ we obviously have $0 \leqslant q < 1$ and hence $|c_k x^k| = |c_k \alpha^k| \cdot \left|\frac{x}{\alpha}\right|^k \leqslant Aq^k$ for all $k = 0, 1, 2, \ldots$. But the series $\sum\limits_{k=0}^{+\infty} Aq^k$ converges as a geometric series with the modulus of its common ratio less than unity and consequently, according to the comparison test (e.g. see [8], Chapter 13, § 2), the series $\sum\limits_{k=0}^{+\infty} |c_k x^k|$ is also convergent. This means that the series $\sum\limits_{k=0}^{+\infty} c_k x^k$ is absolutely convergent for the given x and the lemma has thus been proved.

The lemma implies that if a power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ converges for a value $x = \alpha \neq 0$ it converges absolutely on the interval $-|\alpha| < x < |\alpha|$. In particular, if the series $\sum\limits_{k=0}^{+\infty} c_k x^k$ is convergent on the whole x-axis it is absolutely convergent there.

Proof of Theorem 8.5. Let us put $R = \sup |x'|$ where x' runs through the set of all the points of convergence of the series. We obviously have $R < +\infty$ because, if contrary, there would exist points of convergence x' with arbitrarily large moduli $|x'|$ and hence the series would be absolutely convergent throughout the

---

* As will be shown later, if $(-R, R)$ is the interval of convergence of a power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ the domain of convergence of the series may be the open interval $(-R, R)$ or the closed interval $[-R, R]$ or one of the two half-open intervals $(-R, R]$ and $[-R, R)$.

x-axis (by the above lemma), which contradicts the conditions of the theorem. The definition of the number R suggests that the series diverges for $|x| > R$. Let us prove that it converges absolutely for $|x| < R$. Take an arbitrary x with $|x| < R$. By the definition of the least upper bound of a set of numbers, there is a point of convergence x' such that $|x| < |x'| < R$. But then the lemma indicates that the series is absolutely convergent for this value of x. The theorem has been proved.

The behaviour of a power series at an end point of its interval of convergence is specified by the individual peculiarities of the series. For instance, the series

$$1 + x + x^2 + \ldots + x^n + \ldots \tag{a}$$

$$1 - x + \frac{x^2}{2} - \frac{x^3}{3} + \ldots + (-1)^n \frac{x^n}{n} + \ldots \tag{b}$$

$$1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \ldots + \frac{x^n}{n} + \ldots \tag{c}$$

and

$$1 + x + \frac{x^2}{2^2} + \frac{x^3}{3^2} + \ldots + \frac{x^n}{n^2} + \ldots \tag{d}$$

have the common interval of convergence $-1 < x < 1$. Series (a) is a geometric series whose interval of convergence $-1 < x < 1$ has already been discussed and series (b), (c) and (d) are easily investigated by applying D'Alembert's test. Series (b) ((c)) converges, by Leibniz' test, at the end point $x = 1$ $(x = -1)$ of the interval and diverges at the other end point because it turns into the harmonic series for $x = 1$ $(x = -1)$. Series (d) is convergent (by Cauchy's integral test) at either end point of the interval of convergence.

Thus, if the domain of convergence of a power series $\sum_{k=0}^{+\infty} c_k x^k$ is not either the single point $x = 0$ or the whole x-axis there exists a uniquely specified number R, $0 < R < +\infty$, such that the domain of convergence of the power series is one of the four intervals $(-R, R)$, $(-R, R]$, $[-R, R)$ and $[-R, R]$. The number R is referred to as the **radius of convergence of the power series.**

If a power series is convergent only for $x = 0$ we put, by definition, $R = 0$ and if it converges throughout the x-axis we write $R = +\infty$. This convention makes it possible to apply the notion of radius of convergence to all power series. A power series and, consequently, its radius of convergence are completely specified by the sequence of its coefficients $c_0, c_1, \ldots, c_n, \ldots$. The concluding theorems of this section provide some methods for finding the radius of convergence of a power series from its coefficients.

Theorem 8.6$_1$. *If there exists a limit*

$$\lim_{n\to+\infty}\frac{|c_{n+1}|}{|c_n|}=l, \quad l\geqslant 0$$

(finite or infinite) the radius of convergence of the series $\sum_{k=0}^{+\infty} c_k x^k$ *is equal to* $R=\frac{1}{l}$ $\Big($*the expression* $\frac{1}{l}$ *is understood as being equal to zero for* $l=+\infty$ *and as being equal to* $+\infty$ *if* $l=0\Big)$.

Proof. Applying D'Alembert's test* to the series we find that

$$\lim_{n\to+\infty}\frac{|c_{n+1}|\,|x|^{n+1}}{|c_n|\,|x|^n}=|x|\cdot\lim_{n\to+\infty}\frac{|c_{n+1}|}{|c_n|}=|x|\cdot l$$

If $l=0$ we have $|x|\cdot l=0$ and hence the series converges absolutely for any x in this case, i.e. $R=+\infty$. If $l=+\infty$ and $x\neq 0$ we have $|x|\cdot l=+\infty$ and thus the series diverges for any $x\neq 0$, i.e. $R=0$. Finally, if $0<l<+\infty$ the series diverges in the case $l|x|>1$ and converges absolutely in the case $l|x|<1$ and consequently $R=\frac{1}{l}$. The theorem has been proved.

The following theorem is proved after a manner of Theorem 8.6$_1$ by applying Cauchy's root test.

Theorem 8.6$_2$. *If there is a limit*

$$\lim_{n\to+\infty}\sqrt[n]{|c_n|}=l, \quad l\geqslant 0$$

(finite or infinite) the radius of convergence of the series $\sum_{k=0}^{+\infty} c_k x^k$ *is equal to* $R=\frac{1}{l}$ $\Big($*where the expression* $\frac{1}{l}$ *is again understood as being equal to zero for* $l=+\infty$ *and as being equal to* $+\infty$ *for* $l=0\Big)$.

Theorems 8.6$_1$ and 8.6$_2$ only apply when there is a limit (finite or infinite) $\lim\limits_{n\to+\infty}\frac{|c_{n+1}|}{|c_n|}$ or $\lim\limits_{n\to+\infty}\sqrt[n]{|c_n|}$. The following theorem yields a more general result applicable to any power series.

---

* When D'Alembert's test is applied to a series which is not positive one must take the absolute values of its terms. The same refers to Cauchy's root test used for proving Theorem 8.6

Theorem 8.6₃ (the Cauchy-Hadamard Theorem). The radius of convergence of an arbitrary power series* $\sum_{k=0}^{+\infty} c_k x^k$ *is equal to*

$$R=\frac{1}{l} \quad \text{where} \quad l=\overline{\lim_{n\to+\infty}}\sqrt[n]{|c_n|} \tag{8.69}$$

where the expression $\frac{1}{l}$ *is put equal to zero in the case* $l=+\infty$ *and to* $+\infty$ *in the case* $l=0$.

Note. The symbol $\overline{\lim\limits_{n\to+\infty}}\sqrt[n]{|c_n|}$ denotes the *limit superior* of the sequence of nonnegative numbers $|c_1|$, $\sqrt{|c_2|}$, $\sqrt[3]{|c_3|}$, ..., $\sqrt[n]{|c_n|}$, ... If the sequence is unbounded we put, by definition, $\overline{\lim\limits_{n\to+\infty}}\sqrt[n]{|c_n|}=+\infty$. In case it is bounded the limit superior $\overline{\lim\limits_{n\to+\infty}}\sqrt[n]{|c_n|}$ is equal to the maximum of the abscissas of the limit points of the sequence.

Proof of Theorem 8.6₃. Only the following three cases are possible here: (1) $0<l<+\infty$, (2) $l=0$ and (3) $l=+\infty$. We shall separately consider each case.

(1) Let $0<l<+\infty$. We shall prove that $R=\frac{1}{l}$ which is equivalent to the following two assertions: (a) the series $\sum_{k=0}^{+\infty} c_k x^k$ converges for every x_1 satisfying the condition $|x_1|<\frac{1}{l}$ and (b) the series is divergent for any x_2 such that $|x_2|>\frac{1}{l}$.

(a) Let $|x_1|<\frac{1}{l}$, i.e. $l|x_1|<1$. Then, given a sufficiently small $\varepsilon>0$, we have $(l+\varepsilon)|x_1|=q<1$. The quantity $l=\overline{\lim\limits_{n\to+\infty}}\sqrt[n]{|c_n|}$ being equal to the maximum abscissa of the limit points of the sequence $\{\sqrt[n]{|c_n|}\}$, the inequality $\sqrt[n]{|c_n|}<l+\varepsilon$ holds beginning with a sufficiently large n. Consequently, for all such n we have

$$\sqrt[n]{|c_n|}\,|x_1|<(l+\varepsilon)|x_1|=q<1, \quad \text{i.e.} \quad |c_n|\,|x_1|^n<q^n$$

Therefore, by the comparison test, taking into account that the geometric series $\sum_{n=0}^{+\infty} q^n$ is convergent, we conclude that the series $\sum_{n=0}^{+\infty} |c_n|\,|x_1|^n$ converges which means that the series $\sum_{n=0}^{+\infty} c_n x_1^n$ converges absolutely.

* Hadamard, Jacques Salomon (1865-1963), a noted French mathematician.

(b) Let $|x_2|>\frac{1}{l}$, i.e. $l|x_2|>1$. Then for a sufficiently small ε, $0<\varepsilon<l$, we can write $(l-\varepsilon)|x_2|>1$. The quantity $l=\overline{\lim\limits_{n\to+\infty}}\sqrt[n]{|c_n|}$ being the (maximum) abscissa of a limiting point of the sequence $\{\sqrt[n]{|c_n|}\}$, there is an infinite sequence of indices $n_1<n_2<\ldots<n_k<\ldots$ such that $\sqrt[n_k]{c_{n_k}}>l-\varepsilon$, that is $\sqrt[n_k]{|c_{n_k}|}\,|x_2|>(l-\varepsilon)|x_2|>1$ and $|c_{n_k}|\,|x_2|^{n_k}>1$. Thus, the necessary condition for convergence of a series (requiring that the general term of a convergent series should tend to zero) is violated for the series $\sum\limits_{k=0}^{+\infty}c_kx_2^k$ and hence it diverges.

(2) Let $l=0$. We shall show that in this case $R=+\infty$ which is equivalent to the assertion that the series $\sum\limits_{k=0}^{+\infty}c_kx^k$ is convergent for all x, $-\infty<x<+\infty$. Take a value $x_0\neq 0$. Since $l=0=\overline{\lim\limits_{n\to+\infty}}\sqrt[n]{|c_n|}$ we have, beginning with a sufficiently large n, the inequality $\sqrt[n]{|c_n|}<\frac{1}{2|x_0|}$ which shows that for all such n we can write $\sqrt[n]{|c_n|}\,|x_0|<\frac{1}{2}$ and $|c_n|\,|x_0|^n<\frac{1}{2^n}$. Therefore, by virtue of the comparison test, the series $\sum\limits_{n=0}^{+\infty}|c_n|\,|x_0|^n$ converges and thus the series $\sum\limits_{n=0}^{+\infty}c_nx_0^n$ is absolutely convergent.

(3) Let $l=+\infty$, i.e. let the number sequence $\{\sqrt[n]{|c_n|}\}$ be unbounded. We shall prove that in this case $R=0$ which means that the series $\sum\limits_{n=0}^{+\infty}c_nx_0^n$ diverges for any $x_0\neq 0$. Suppose that the series $\sum\limits_{n=0}^{+\infty}c_nx_0^n$ is convergent for a value $x_0\neq 0$. Then, by the above mentioned necessary condition for convergence of a series, we have $c_nx_0^n\to 0$ as $n\to+\infty$ and, consequently, there is a number A, $1<A<+\infty$, such that $|c_nx_0^n|<A$ for all $n=0, 1, 2, \ldots$. Therefore the inequality $\sqrt[n]{|c_n|}\,|x_0|<\sqrt[n]{A}\leqslant A$ holds for all $n=0, 1, 2, \ldots$ and hence $\sqrt[n]{|c_n|}<\frac{A}{|x_0|}$ which contradicts the hypothesis that the sequence $\{\sqrt[n]{|c_n|}\}$ is unbounded. The proof of the theorem has thus been completed.

Note. The Cauchy-Hadamard theorem provides an approach (other than the one applied in Sec. 3) to proving the possibility

of term-by-term differentiation and integration of a power series because, on the basis of formula (8.69), it can be easily verified that the series obtained from a given power series by integration or differentiation has the same radius of convergence as the original series.

2. On Uniform Convergence of a Power Series and Continuity of Its Sum. We have established that every power series is absolutely convergent at the interior points of its interval of convergence. We now proceed to investigate the properties of the power series concerning their uniform convergence.

These properties are described by the following

Theorem 8.7. Every power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ *converges uniformly on each closed subinterval strictly contained within its interval of convergence.*

Proof. Let $-R < \alpha \leqslant x \leqslant \beta < R$ where $(-R, R)$ is the interval of convergence. We shall prove that the series is uniformly convergent on the closed interval $[\alpha, \beta]$. Take a value $x_0 > \max(|\alpha|, |\beta|)$, $x_0 \in (-R, R)$. Then we have the relation $|x| < |x_0|$ for all $x \in [\alpha, \beta]$. This implies the inequality $|c_n x^n| \leqslant |c_n x_0^n|$. But the number series $\sum\limits_{n=0}^{\infty} c_n x_0^n$ is convergent and consequently, by Weierstrass' test, the series $\sum\limits_{n=0}^{+\infty} c_n x^n$ and $\sum\limits_{n=0}^{+\infty} |c_n x^n|$ converge uniformly on the interval $[\alpha, \beta]$. The theorem has been proved.

Note. A power series may not be uniformly convergent on the entire interval of convergence. For example, the series

$$\frac{1}{1-x} = 1 + x + x^2 + \ldots + x^n + \ldots$$

converges nonuniformly on its interval of convergence $-1 < x < 1$ because the modulus of the difference between its sum and its nth partial sum $\frac{1}{1-x} - 1 - x - x^2 - \ldots - x^n$ can be made, for any fixed n, to be arbitrarily large as $x \to 1 - 0$, and consequently, the absolute value of the difference cannot remain smaller than a finite number $\varepsilon > 0$ for all x belonging to the interval $-1 < x < 1$.

On the other hand, the following theorem takes place:

Theorem 8.7₁. If a power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ *converges at the end point $x = R$ of its interval of convergence $(-R, R)$ it is uniformly convergent on the closed interval $[0, R]$.*

Proof. We shall show that under the assumption of the theorem the condition of Cauchy's test for uniform convergence is fulfilled on the closed interval $[0, R]$. This will imply the uniform convergence of the series on $[0, R]$.

Let us introduce the notation

$$S_{n,\,p}=c_{n+1}R^{n+1}+\ldots+c_{n+p}R^{n+p},\quad p=1,\,2,\,\ldots$$

We obviously have

$$c_{n+1}R^{n+1}=S_{n,\,1}$$

$$c_{n+2}R^{n+2}=S_{n,\,2}-S_{n,\,1},\;\ldots,\;c_{n+p}R^{n+p}=S_{n,\,p}-S_{n,\,p-1} \tag{α}$$

Let there be given $\varepsilon>0$. The number series $\sum\limits_{k=0}^{+\infty}c_kR^k$ being convergent (by the hypothesis), Cauchy's test for a convergent number series suggests that there is $N(\varepsilon)$ such that we have

$$|S_{n,\,k}|<\varepsilon\quad\text{for all } k=0,\,1,\,2,\,3,\,\ldots \tag{β}$$

when $n>N(\varepsilon)$. Taking into account that $\left(\frac{x}{R}\right)^{n+p}\leqslant\left(\frac{x}{R}\right)^{n+p-1}\leqslant\ldots$ $\ldots\leqslant\left(\frac{x}{R}\right)^{n}\leqslant1$ for $0\leqslant x\leqslant R$ and making use of (α) and (β) we obtain

$$|c_{n+1}x^{n+1}+c_{n+2}x^{n+2}+\ldots+c_{n+p}x^{n+p}|=$$

$$=\left|c_{n+1}R^{n+1}\left(\frac{x}{R}\right)^{n+1}+c_{n+2}R^{n+2}\left(\frac{x}{R}\right)^{n+2}+\right.$$

$$\left.\ldots+c_{n+p}R^{n+p}\left(\frac{x}{R}\right)^{n+p}\right|=\left|S_{n,\,1}\left[\left(\frac{x}{R}\right)^{n+1}-\left(\frac{x}{R}\right)^{n+2}\right]+\right.$$

$$+S_{n,\,2}\left[\left(\frac{x}{R}\right)^{n+2}-\left(\frac{x}{R}\right)^{n+3}\right]+\ldots$$

$$\left.\ldots+S_{n,\,p-1}\left[\left(\frac{x}{R}\right)^{n+p-1}-\left(\frac{x}{R}\right)^{n+p}\right]+S_{n,\,p}\left(\frac{x}{R}\right)^{n+p}\right|\leqslant$$

$$\leqslant|S_{n,\,1}|\left[\left(\frac{x}{R}\right)^{n+1}-\left(\frac{x}{R}\right)^{n+2}\right]+|S_{n,\,2}|\left[\left(\frac{x}{R}\right)^{n+2}-\left(\frac{x}{R}\right)^{n+3}\right]+\ldots$$

$$\ldots+|S_{n,\,p-1}|\left[\left(\frac{x}{R}\right)^{n+p-1}-\left(\frac{x}{R}\right)^{n+p}\right]+|S_{n,\,p}|\left(\frac{x}{R}\right)^{n+p}<$$

$$<\varepsilon\left\{\left(\frac{x}{R}\right)^{n+1}-\left(\frac{x}{R}\right)^{n+2}+\left(\frac{x}{R}\right)^{n+2}-\left(\frac{x}{R}\right)^{n+3}+\ldots\right.$$

$$\left.\ldots-\left(\frac{x}{R}\right)^{n+p}+\left(\frac{x}{R}\right)^{n+p}\right\}=\varepsilon\left(\frac{x}{R}\right)^{n+1}\leqslant\varepsilon$$

for all $n>N(\varepsilon)$, all $p=1,\,2,\,\ldots$ and simultaneously for all x belonging to the interval $0\leqslant x\leqslant R$, which is what we set out to prove.

Note 1. We have a similar situation in case a series of the form $\sum\limits_{k=0}^{+\infty} c_k x^k$ converges at the left end point of its interval of convergence $(-R, R)$ or at both end points: in the former case the series is uniformly convergent on $[-R, 0]$ and in the latter it uniformly converges on $[-R, R]$.

Note 2. If a power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ diverges at the end point $x = R$ of its interval of convergence $(-R, R)$ it cannot be uniformly convergent on the interval $0 \leqslant x < R$. For, assuming the contrary, we would conclude, on the basis of the theorem on passing to limit in a uniformly convergent series, that the series $\sum\limits_{k=0}^{+\infty} c_k x^k$ converges at the end point $x = R$ as well, which contradicts the original condition.

The theorem below is a consequence of the theorems on uniform convergence of a power series and of the fact that the terms of a power series are continuous functions.

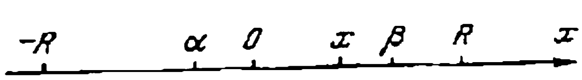
Theorem 8.8. The sum of a power series

$$S(x) = \sum_{k=0}^{+\infty} c_k x^k \tag{8.70}$$

*is continuous at every interior point of its interval of convergence.**

Proof. If a point x lies in the interior of the interval of convergence $(-R, R)$ of series (8.70) the point can be embedded in a closed

Fig. 8.4



interval $[\alpha, \beta]$, $-R < \alpha < \beta < R$, strictly contained within the interval of convergence (Fig. 8.4). Series (8.70) converging uniformly on the interval $[\alpha, \beta]$, its sum is a continuous function on this interval because the terms of the series are continuous on $[\alpha, \beta]$. Hence, the sum $S(x)$ is continuous at the point $x \in [\alpha, \beta]$. The theorem has been proved.

Note. If series (8.70) converges at an end point of its interval of convergence $(-R, R)$ its sum is continuous at the end point. This follows from the uniform convergence of series (8.70) on the corresponding closed interval of the form $[-R, 0]$ or $[0, R]$ (see Theorem 8.7_1).

---

* Here and henceforward we assume that the interval of convergence of series (8.70) does not degenerate into the point $x = 0$.

3. Differentiation and Integration of Power Series.

Theorem 8.9. Power series (8.70) *admits term-by-term differentiation at the interior points of its interval of convergence, that is its sum* $S(x)$ *is differentiable, the relation*

$$S'(x)=\sum_{k=1}^{+\infty} kc_k x^{k-1} \tag{8.71}$$

holds for each $x \in (-R, R)$ *and differentiated series* (8.71) *has the same interval of convergence.*

Proof. Let R and R' designate, respectively, the radii of convergence of series (8.70) and (8.71). Let us first prove that $R' = R$. If $x \in (-R', R')$ the series $\sum_{k=1}^{+\infty} k\,|c_k|\,|x|^{k-1}$ is convergent and therefore the series $\sum_{k=1}^{+\infty} k\,|c_k|\,|x|^{k}$ is convergent as well. Consequently, by the comparison test, the series $\sum_{k=0}^{+\infty} |c_k|\,|x|^{k}$ also converges and hence $x \in (-R, R)$. Therefore $R' \leqslant R$. Furthermore, if $x \in (-R, R)$ we can choose $x_0 \in (-R, R)$ such that the inequality $|x_0| > |x|$ $(x_0 \neq 0)$ is fulfilled. The series $\sum_{k=0}^{+\infty} c_k x_0^k$ converging, we have $c_k x_0^k \to 0$ as $k \to +\infty$. Hence, there is a constant $A > 0$ such that $|c_k x_0^k| < A$ for all $k = 0, 1, 2, \ldots$. This enables us to estimate the terms of series (8.71) as follows:

$$|kc_k x^{k-1}| = \frac{1}{|x_0|}k|c_k x_0^k|\cdot\left|\frac{x}{x_0}\right|^{k-1} < \frac{A}{|x_0|}k\left|\frac{x}{x_0}\right|^{k-1} \tag{8.72}$$

For $|x| < |x_0|$, i.e. $\left|\frac{x}{x_0}\right| < 1$, the series $\sum_{k=1}^{+\infty}\frac{A}{|x_0|}k\left|\frac{x}{x_0}\right|^{k-1}$ converges, which can be easily established on the basis of D'Alembert's test. Indeed, for $k \to +\infty$ we have

$$\frac{A}{|x_0|}(k+1)\left|\frac{x}{x_0}\right|^{k+1} : \frac{A}{|x_0|}k\left|\frac{x}{x_0}\right|^{k} = \left(1+\frac{1}{k}\right)\left|\frac{x}{x_0}\right| \to \left|\frac{x}{x_0}\right| < 1$$

Now applying the comparison test we conclude, by relation (8.72), that series (8.71) also converges for the given x, i.e. $x \in (-R', R')$. Consequently, $R \leqslant R'$. This result (together with the above inequality $R' \leqslant R$) implies that $R' = R$.

Now we can make use of the fact that series (8.70) and (8.71) with continuous terms are uniformly convergent on every closed subinterval strictly contained in the interior of their common interval of convergence $(-R, R)$, which indicates that the conditions of Theo-

rem 8.3 on termwise differentiation of a functional series are fulfilled whence it follows that Theorem 8.9 is true.

Corollary. The sum of a power series $S(x)=\sum_{k=0}^{+\infty} c_k x^k$ *possesses the derivatives of all orders and*

$$S^{(n)}(x)=\sum_{k=n}^{+\infty} k(k-1)\ldots(k-n+1)\,c_k x^{k-n},\quad n=1,\ 2,\ \ldots \tag{8.73}$$

Besides, the radius of convergence of series (8.73) *coincides with that of series* $S(x)=\sum_{k=0}^{+\infty} c_k x^k$.

The proof of the corollary is based on Theorem 8.9 which is consecutively applied to differentiated series (8.71), then to the series thus obtained and so on.

Theorem 8.10. *It is permissible to integrate a power series*

$$S(x)=\sum_{k=0}^{+\infty} c_k x^k \tag{8.70}$$

termwise between any limits of integration lying within its interval of convergence. In particular, we have the relation

$$\int_0^x S(z)\,dz=\sum_{k=0}^{+\infty} c_k \frac{x^{k+1}}{k+1},\qquad x\in(-R,\ R) \tag{8.74}$$

Besides, the radii of convergence of the integrated series and series (8.70) *coincide.*

Proof. The terms of series (8.70) are continuous functions and hence the series can be integrated term-by-term on every interval of uniform convergence. Given an arbitrary point $x\in(-R,\ R)$, we can embed it in a closed interval $[\alpha,\ \beta]$ strictly contained in the interval $(-R,\ R)$. Furthermore, such an interval $[\alpha,\ \beta]$ can always be constructed so that it should contain the origin of coordinates. Hence, series (8.70) can in fact be integrated termwise. In particular, integrating series (8.70) from 0 to x on the corresponding interval of type $[\alpha,\ \beta]\subset(-R,\ R)$ we derive equality (8.74) and hence the proof has been completed.

Note. If series (8.70) converges at an end point of its interval of convergence $(-R,\ R)$ the limit of integration x in equality (8.74) can be made to coincide with that end point because in this case series (8.70) is uniformly convergent on the corresponding closed interval $[-R,\ 0]$ or $[0,\ R]$.

4. Arithmetical Operations on Power Series. Let us first study the operations of addition, subtraction and multiplication. Consider

two series

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots = \sum_{n=0}^{+\infty} a_nx^n \qquad (\alpha)$$

and

$$g(x) = b_0 + b_1x + b_2x^2 + \ldots + b_nx^n + \ldots = \sum_{n=0}^{+\infty} b_nx^n \qquad (\beta)$$

whose radii of convergence are, respectively, equal to $R_a > 0$ and $R_b > 0$. Then

$$f(x) \pm g(x) = \sum_{n=0}^{+\infty} (a_n \pm b_n)\, x^n \quad \text{for} \quad |x| < \min(R_a, R_b) \qquad (\gamma)$$

and

$$f(x)\, g(x) = \sum_{n=0}^{+\infty} (a_0b_n + a_1b_{n-1} + \ldots + a_nb_0)\, x^n \text{ for } |x| < \min(R_a, R_b) \quad (\delta)$$

Relation (γ) is an apparent consequence of the corresponding theorem on addition and subtraction of number series (e.g. see [8], Chapter 13, § 4). Relation (δ) follows from the theorem on multiplication of absolutely convergent number series (e.g. see [8], Chapter 13, § 4) since both series (α) and (β) are absolutely convergent for $|x| < \min(R_a, R_b)$.

Finally, let us consider the operation of division. If $R_a > 0$, $R_b > 0$ and $b_0 \neq 0$ we can write, for sufficiently small values of $|x|$, a power series expansion of the quotient of series (α) and (β):

$$\frac{a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots}{b_0 + b_1x + b_2x^2 + \ldots + b_nx^n + \ldots} = c_0 + c_1x + c_2x^2 + \ldots + c_nx^n + \ldots$$

The coefficients $c_0, c_1, \ldots, c_n, \ldots$ of the expansion can be successively found by means of the recurrence formulas which are obtained if we perform the multiplication of the power series on the right-hand side of the identity

$$a_0 + a_1x + a_2x^2 + \ldots \equiv$$
$$\equiv (b_0 + b_1x + b_2x^2 + \ldots)(c_0 + c_1x + c_2x^2 + \ldots)$$

and compare the coefficients in like powers of x on the left-hand and right-hand sides of the resulting relation. The series $\sum_{n=0}^{+\infty} c_nx^n$ can also be obtained by dividing the series $a_0 + a_1x + a_2x^2 + \ldots$ by the series $b_0 + b_1x + b_2x^2 + \ldots$ according to the same rule as those applied in dividing polynomials arranged in ascending powers of the variable. We shall not present here the proofs of these assertions.

§ 4. EXPANDING FUNCTIONS IN POWER SERIES

We say that a function $f(x)$ can be *expanded in a power series* $\sum_{k=0} c_k x^k$ *on an interval* $(-r, r)$ if the series is convergent on the interval and its sum is equal to $f(x)$, i.e.

$$f(x) = \sum_{k=0}^{+\infty} c_k x^k \tag{8.75}$$

on the interval $(-r, r)$.* We suppose, of course, that the interval $(-r, r)$ does not degenerate into a point. We have already discussed the importance of expansions of functions in power series and functional series of other types (see the beginning of the present chapter). At the end of § 4 we shall present some characteristic examples of applications of power series expansions for real values of the variable x. In § 5 we shall consider some basic properties of power series in a complex variable.

1. Key Theorems on Expanding Functions in Power Series. Expanding Elementary Functions. First of all we prove that a function $f(x)$ cannot have two different expansions of form (8.75):

Theorem 8.11. A power series

$$f(x) = \sum_{k=0}^{+\infty} c_k x^k \tag{8.76}$$

convergent on an interval $(-R, R)$ *(which does not degenerate into a point) is Taylor's** series for its sum* $f(x)$, *that is the one whose coefficients are determined by Taylor's formula*

$$c_k = \frac{f^{(k)}(0)}{k!}, \quad k = 0, 1, 2, \ldots \tag{8.77}$$

Hence the coefficients of power series (8.76) *are uniquely specified by its sum.*

Proof. To prove the theorem we can make use of the corollary of the theorem on termwise differentiation of a power series. It follows that the sum $f(x)$ of series (8.75) is infinitely differentiable and the equality

$$f^{(n)}(x) = \sum_{k=n}^{+\infty} k(k-1)\ldots(k-n+1)\, c_k x^{k-n}, \qquad n = 1, 2, \ldots \tag{8.78}$$

---

* A function $f(x)$ which can be expanded in a power series on an interval $(-r, r)$ is said to be an analytic function of the variable x on that interval.

** Taylor, Brook (1685-1731), an English mathematician.

takes place. Putting $x = 0$ in (8.78) we obtain

$$f^{(n)}(0) = n!c_n$$

and, consequently,

$$c_n = \frac{f^{(n)}(0)}{n!}$$

which is what we set out to prove. Thus, if a function $f(x)$ can be expanded into a power series convergent to the function this series is Taylor's series for the function.

Now it appears natural to pose the question whether the converse assertion is true. The problem can be stated as follows. Suppose a function $f(x)$ is infinitely differentiable on an interval $(-R, R)$ where $R \neq 0$. We can formally construct the Taylor series for this function:

$$f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \ldots \tag{8.79}$$

Now, does series (8.79) converge on the interval $(-R, R)$ and will its sum be equal to the function $f(x)$ in case it does? It turns out that in the general case the answer to the question is negative which can be confirmed by the example of the function

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases} \tag{8.80}$$

In fact, it can be easily verified that the function is infinitely differentiable throughout the x-axis and that at the origin we have

$$f(0) = f'(0) = \ldots = f^{(n)}(0) = f^{(n+1)}(0) = \ldots = 0 \tag{8.81}$$

Consequently, all the coefficients of the Taylor series of the function are equal to zero. Thus, the Taylor series converges on the entire x-axis and its sum is identically equal to zero whereas the function takes on a zero value only at the origin.

Theorem 8.12. A function $f(x)$ can be expanded in a power series $\sum\limits_{k=0}^{+\infty} c_k x^k$ on an interval $(-R, R)$ if and only if the function $f(x)$ possesses the derivatives of all orders on the interval and R_n, the ***remainder after n terms R_n in Taylor's formula,***

$$f(x) = f(0) + f'(0)x + \ldots + \frac{f^{(n)}(0)}{n!}x^n + R_n \tag{8.82}$$

tends to zero for every fixed $x \in (-R, R)$ as $n \to \infty$.

Proof. If a function $f(x)$ can be expanded in a power series $\sum\limits_{k=0}^{\infty} c_k x^k$ on an interval $(-R, R)$ it follows from the corollary of Theorem 8.9

that $f(x)$ has the derivatives of all orders and hence, by Theorem 8.11, the equality $f(x) = \sum_{k=0}^{+\infty} c_k x^k$ can be rewritten in the form

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \ldots \tag{8.83}$$

Equality (8.83) suggests that the difference between the sum $f(x)$ and the nth partial sum of series (8.83) tends to zero, as $n \to +\infty$, for all $x \in (-R, R)$. The difference being, by definition, the remainder R_n of Taylor's formula, the necessity has thus been proved.

Conversely, if $f(x)$ possesses the derivatives of all orders on the interval $(-R, R)$ and if R_n (the remainder in Taylor's formula (8.82)) tends to zero as $n \to +\infty$ for every $x \in (-R, R)$, we have

$$\left| f(x) - \left[f(0) + \frac{f'(0)}{1!}x + \ldots + \frac{f^{(n)}(0)}{n!}x^n \right] \right| \to 0$$

for $n \to +\infty$ and each $x \in (-R, R)$. Consequently, series (8.83) converges on the interval $(-R, R)$ and its sum is equal to $f(x)$, which is what we set out to prove.

The following theorem yields some convenient sufficient conditions for a function to have a power series expansion.

Theorem 8.13. If a function $f(x)$ has the derivatives of all orders on an interval $(-R, R)$ and if there exists a positive constant M such that

$$|f^{(n)}(x)| \leqslant M \text{ for } n = 0, 1, 2, \ldots \text{ and all } x \in (-R, R) \tag{8.84}$$

that is if the family of the derivatives of all orders is uniformly bounded on the interval $(-R, R)$, the function $f(x)$ can be expanded into a power series on $(-R, R)$.

Proof. The derivatives of all orders existing for the function $f(x)$ on the interval $(-R, R)$, we can formally construct its Taylor series. Let us prove that the series converges to $f(x)$. According to Theorem 8.12, for this purpose it is sufficient to show that the remainder in Taylor's formula (8.82) tends to zero, as $n \to +\infty$, for all $x \in (-R, R)$. Applying *Lagrange's form of the remainder for Taylor's theorem** we derive, on the basis of (8.84), the following inequality for R_n:

$$|R_n| = \left| \frac{f^{(n+1)}(\theta x)}{(n+1)!} x^{n+1} \right| < \frac{MR^{n+1}}{(n+1)!} \quad \text{for} \quad n = 0, 1, \ldots, x \in (-R, R)$$
$$(0 < \theta < 1) \tag{8.85}$$

---

* Lagrange, Joseph Louis (1736-1813), a famous French mathematician. For Lagrange's form of the remainder for Taylor's theorem see, for instance, [8] Chapter 8, § 9.

Making use of D'Alembert's test we can easily verify that the series $\sum_{n=0}^{+\infty} \frac{MR^{n+1}}{(n+1)!}$ is convergent. Therefore, by the necessary condition for convergence of a series, we have $\frac{MR^{n+1}}{(n+1)!} \to 0$ for $n \to +\infty$. Hence, by virtue of inequality (8.85), R_n tends to zero, as $n \to +\infty$, for all $x \in (-R, R)$, which is what we set out to prove.

Let us now proceed to consider some important examples of expanding functions in power series, i.e. in Taylor's series.

1. For the functions $f(x) = \sin x$ and $f(x) = \cos x$ we have $|f^{(n)}(x)| \leqslant 1$ for all $n = 0, 1, 2, \ldots$ and all x, $-\infty < x < +\infty$. Therefore either function can be expanded in its Taylor series convergent throughout the number axis.

Computing Taylor's coefficients $\frac{f^{(n)}(0)}{n!}$ we obtain

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \ldots$$

and

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots$$

2. The derivatives of all orders of the function $f(x) = e^x$ satisfy the inequality $|f^{(n)}(x)| = e^x \leqslant e^R$ on an arbitrary interval $(-R, R)$. Consequently, the exponential function $f(x) = e^x$ is expanded in a power series on every interval $(-R, R)$ of the x-axis, that is on the entire x-axis.

Computing Taylor's coefficients we find

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots$$

3. To investigate the power series expansion of the function $f(x) = \ln(1 + x)$ it is advisable to apply the following technique. Differentiating with respect to x and expanding the derivative thus obtained in a geometric series we find

$$f'(x) = \frac{1}{1+x} = 1 - x + x^2 - x^3 + \ldots \quad \text{for } -1 < x < 1$$

Now, integrating this equality term-by-term we receive, on the basis of Theorem 8.3, the formula

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \quad \text{for } -1 < x < 1 \qquad (8.86)$$

Expansion (8.86) also remains valid for $x = 1$. Indeed, by Leibniz' test, series (8.86) converges for $x = 1$ and therefore its sum

$$S(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$$

is continuous (see the note after Theorem 8.8) on the closed interval $[0, 1]$. Consequently,

$$\lim_{x\to 1-0} S(x) = S(1) = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$$

The function $f(x) = \ln(1+x)$ is also continuous on that interval and therefore $\ln 2 = \ln(1+x)\,|_{x=1-0}$. But, according to (8.86), we have the equality $\ln(1+x) = S(x)$ for $0 \leqslant x < 1$. Hence,

$$\ln 2 = \lim_{x\to 1-0} \ln(1+x) = \lim_{x\to 1-0} S(x) = S(1) = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$$

4. The function $f(x) = \arctan x$ can be expanded in a power series by applying the same procedure. Thus, differentiating and expanding the derivative in a geometric series we get

$$f'(x) = \frac{1}{1+x^2} = 1-x^2+x^4-x^6+\ldots \quad \text{for} \quad -1<x<1$$

Performing termwise integration we find

$$\arctan x = x-\frac{x^3}{3}+\frac{x^5}{5}-\frac{x^7}{7}+\ldots \quad \text{for} \quad -1<x<1$$

The validity of this expansion for the point $x = 1$ can be proved after a manner of the above proof of the validity of the expansion

$$\ln 2 = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$$

5. Let us formally write Taylor's formula for the function $f(x) = (1+x)^\alpha$ where α is an arbitrary real number distinct from zero:

$$1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+$$

$$+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3+\ldots = 1+\sum_{k=1}^{+\infty}\frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!}x^k \quad (8.87)$$

For a positive integral $\alpha = n$ all the terms of series (8.87) from the $(n+1)$th onwards turn into zero and thus we arrive at the well known *Newton binomial formula*. Applying D'Alembert's test to the case $\alpha \neq n$, $n = 0, 1, 2, \ldots$, we can easily show that the radius of convergence of series (8.87) is equal to unity. Consequently, outside the closed interval $[-1, 1]$, the function $(1+x)^\alpha$ (for α unequal to a positive integer) cannot be expanded into a power series of the form $\sum\limits_{k=0}^{+\infty} c_k x^k$, i.e. a series in positive integral powers of x. Let us prove that inside the interval the expansion

$$(1+x)^\alpha = 1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\ldots = 1+\sum_{k=1}^{+\infty}\frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!}x^k \quad (8.88)$$

is valid. According to Theorem 8.12, for this purpose it will suffice to show that the remainder in Taylor's formula for the function $(1+x)^{\alpha}$ tends to zero as $n \to +\infty$ for all x belonging to the interval $(-1, 1)$.

We shall make use of *Cauchy's form of the remainder for Taylor's theorem*:

$$R_n = \frac{(1-\theta)^n x^{n+1}}{n!} f^{(n+1)}(\theta x), \quad 0<\theta<1$$

For the function $(1+x)^{\alpha}$ this results in

$$R_n = \frac{(1-\theta)^n x^{n+1}}{n!} \alpha(\alpha-1) \ldots (\alpha-n)(1+\theta x)^{\alpha-n} =$$

$$= \left[\frac{(\alpha-1)(\alpha-2)\ldots(\alpha-n)}{n!} x^n\right] \cdot \left[\left(\frac{1-\theta}{1+\theta x}\right)^n\right] \cdot \left[\alpha x(1+\theta x)^{\alpha}\right] \quad (8.89)$$

We consider the values $x>-1$ and thus $0<\frac{1-\theta}{1+\theta x}<1$. Consequently, we have $0<\left(\frac{1-\theta}{1+\theta x}\right)^n<1$ for all $n=1, 2, 3, \ldots$. Furthermore, the quantity $|\alpha x|(1+\theta x)^{\alpha}$ lies between the limits $|\alpha x|(1-|x|)^{\alpha}$ and $|\alpha x|(1+|x|)^{\alpha}$ (independent of n) for all $x \in (-1, 1)$ because we have the inequality $0<\theta<1$. Finally, the factor $\frac{(\alpha-1)(\alpha-2)\ldots(\alpha-n)}{n!} x^n$ is the nth term of Taylor's series for the function $(1+x)^{\alpha-1}$ whose convergence for $-1<x<1$ is easily established by D'Alembert's test. Therefore, for $-1<x<1$, we conclude, by the necessary condition for convergence of a series, that

$$\frac{\alpha(\alpha-1)\ldots(\alpha-n)}{n!} x^n \to 0$$

as $n \to +\infty$. Thus, the second and the third factors in the square brackets on the right-hand side of relation (8.89) are bounded for $-1<x<1$ whereas the first factor tends to zero as $n \to +\infty$. Hence, $R_n \to 0$ for every $x \in (-1, 1)$ as $n \to +\infty$. This completes the proof of the validity of expansion (8.88) on the interval $(-1, 1)$.

2. Some Applications of Power Series.

(a) *Applying Power Series to Computing Approximate Values of Functions.* The series

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots \quad (8.90)$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^k \frac{x^{2k+1}}{(2k+1)!} + \cdots \quad (8.91)$$

and

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + (-1)^k \frac{x^{2k}}{(2k)!} + \cdots \quad (8.92)$$

can be used for computing the values of e^x, $\sin x$ and $\cos x$ for any values of x with an arbitrary accuracy since relations (8.90)-(8.92) hold throughout the x-axis for these functions.

Taking partial sums of series (8.90)-(8.92) as approximate values of the corresponding functions we can easily estimate the error because, according to Leibniz' test*, in the case of series (8.91) and (8.92) the error does not exceed (in its absolute value) the modulus of the first discarded term.

Although the series

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^{k-1}\frac{x^k}{k!} + \ldots \qquad (8.93)$$
$$-1 < x \leqslant 1$$

for the logarithmic function is an alternating one it converges very slowly. For the values $x > 1$ the series is divergent. To accelerate the convergence of the series and apply it to computing logarithms of numbers greater than unity we subtract the expansion

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \ldots \qquad (8.93')$$

from expansion (8.93). This results in

$$\ln\left(\frac{1+x}{1-x}\right) = 2x\left(1 + \frac{x^2}{3} + \frac{x^4}{5} + \ldots\right) \qquad (8.94)$$

Putting $x = \frac{1}{2n+1}$ in (8.94) we obtain the formula

$$\ln\frac{n+1}{n} = \ln(n+1) - \ln n =$$
$$= \frac{2}{2n+1}\left(1 + \frac{1}{3(2n+1)^2} + \frac{1}{5(2n+1)^4} + \ldots\right) \qquad (8.95)$$

Series (8.95) converges sufficiently fast, which makes it possible to apply it to computing logarithms of natural numbers starting from the value $\ln 1 = 0$.

The series for the arctangent

$$\operatorname{arc\,tan} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots \qquad (8.96)$$

can be used for computing the number π with an arbitrary accuracy. Namely, putting $x = 1$ in (8.96) we obtain

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots \qquad (8.97)$$

Since (8.97) is an alternating series, the error arising when its sum is replaced by a partial sum can be easily estimated.

---

* E.g. see [8], Chapter 13, § 5

The series

$$(1+x)^{\alpha}=1+\sum_{k=1}^{+\infty}\frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}x^{k}$$

can be applied to extracting roots. For instance,

$$\sqrt[3]{10}=2\sqrt[3]{1+\frac{1}{4}}=2\left(1+\frac{1}{4}\right)^{\frac{1}{3}}=2\left[1+\frac{1}{3}\cdot\frac{1}{4}-\frac{2}{9}\left(\frac{1}{4}\right)^{2}+\dots\right] \tag{8.98}$$

Since (8.98) is an alternating series, the error of an approximation can be easily estimated. In particular, the sum of the terms written in full on the right-hand side of (8.98) yields an approximation of the root correct to the nearest ten thousandth.

(b) ***Applying Power Series Expansions to Computing Integrals Inexpressible in Terms of Elementary Functions.*** As an example of such an application let us use series expansion (8.91) of the sine for computing the integral $\mathrm{Si}\,x=\int\limits_0^x \frac{\sin\xi}{\xi}\,d\xi$ (the *sine integral*):

$$\mathrm{Si}\,x=\int\limits_0^x\frac{\sin\xi}{\xi}\,d\xi=x-\frac{x^3}{3\cdot 3!}+\frac{x^5}{5\cdot 5!}-\frac{x^7}{7\cdot 7}+\dots \tag{8.99}$$

It should be noted that the division of series (8.91) by x is permissible here for $x\neq 0$ and therefore we have

$$\frac{\sin x}{x}=1-\frac{x^2}{3!}+\frac{x^4}{5!}-\dots,\quad x\neq 0 \tag{*}$$

Putting $\frac{\sin x}{x}=1$ at the point $x=0$ we can retain equality (*) for $x=0$. The expression on the right-hand side of (8.99) is an alternating series and hence the error of an approximation obtained by replacing its sum by a partial sum can be easily estimated.

(c) ***Applying Power Series to Integrating Differential Equations.*** Power series and series in fractional powers of an argument are widely applied for constructing solutions of differential equations and introducing new classes of functions. For greater detail concerning these questions the reader is referred to courses on differential equations, e.g. [5]. Here we shall restrict ourselves to an elementary example. Suppose it is necessary to expand into a power series the function $F(x)=e^{-x^2}\int\limits_0^x e^{\xi^2}\,d\xi$. It can be

easily verified that $F(x)$ satisfies the differential equation

$$F'(x) + 2xF(x) = 1 \tag{8.100}$$

with the initial condition $F(0) = 0$. Let us try to find the solution of equation (8.100) satisfying the condition $F(0) = 0$ in the form of a power series

$$F(x) = \sum_{n=0}^{+\infty} c_n x^n = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n + \ldots \tag{8.101}$$

Substituting the series into equation (8.100) and equating the coefficients in the same powers of x on the left-hand and right-hand sides of the resulting equality we obtain

$$c_1 = 1, \quad (n+2)\, c_{n+2} + 2c_n = 0, \quad n = 1, 2, \ldots \tag{8.102}$$

The initial condition $F(0) = 0$ implies

$$c_0 = 0 \tag{8.103}$$

By (8.102) and (8.103), it follows from (8.101) that

$$F(x) = e^{-x^2} \int_0^x e^{\xi^2}\, d\xi = \sum_{n=0}^{+\infty} \frac{(-1)^n 2^n x^{2n+1}}{1 \cdot 3 \ldots (2n+1)} \tag{8.104}$$

Here we have first performed formal term-by-term differentiation of series (8.101) but now, as its coefficients are known, we see that series (8.104) converges for all x, $-\infty < x < +\infty$, and consequently the termwise differentiation is in fact permissible for all the values of x (see Theorem 8.9).

§ 5. POWER SERIES IN COMPLEX ARGUMENT

A sequence of complex numbers

$$z_1 = x_1 + iy_1, \quad z_2 = x_2 + iy_2, \quad \ldots, \quad z_n = x_n + iy_n, \quad \ldots \tag{A}$$

is said to be *convergent* to a complex number $z_0 = x_0 + iy_0$ as $n \to +\infty$ if

$$|z_n - z_0| = \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2} \to 0 \quad \text{for } n \to +\infty$$

Therefore, for sequence (A) to converge to a number $z_0 = x_0 + iy_0$ it is necessary and sufficient that

$$x_n \to x_0 \quad \text{and} \quad y_n \to y_0 \quad \text{for} \quad n \to +\infty$$

A series with complex terms

$$\sum_{k=1}^{+\infty} a_k$$

where $a_k = \alpha_k + i\beta_k$, $k = 1, 2, \ldots$, is called *convergent* if the sequence of its partial sums converges.

The notions of absolute and conditional convergence as well as Cauchy's necessary and sufficient condition for convergence, D'Alembert's test and Cauchy's root test are easily extended to the series with complex terms.*

The domain of convergence of a power series

$$c_0 + c_1 z + c_2 z^2 + \ldots + c_n z^n + \ldots \qquad \text{(B)}$$

where $c_0, c_1, \ldots, c_n, \ldots$ are complex numbers and $z = x + iy$ is a complex variable is a circle **(the circle of convergence)** with centre at the point $z = 0$. The radius of this circle is **the radius of convergence** of series (B). The circle may degenerate into the point $z = 0$ or expand over the whole plane of the variable $z = x + iy$. In the interior of the circle of convergence series (B) is absolutely convergent. These assertions follow from the lemma below.

Lemma. If power series (B) *converges for $z = \alpha \neq 0$ it is absolutely convergent for every z satisfying the inequality $|z| < |\alpha|$, i.e. in the circle of radius $|\alpha|$ with centre at the point $z = 0$.***

In every circle concentric with the circle of convergence and strictly contained in it the power series converges uniformly and its sum is not only continuous but also infinitely differentiable.

Proceeding from the power series expansions of the elementary functions of a real variable

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots \qquad (8.105)$$

$$\sin x = x - \frac{x^3}{3!} + \ldots + (-1)^k \frac{x^{2k+1}}{(2k+1)!} + \ldots \qquad (8.106)$$

and

$$\cos x = 1 - \frac{x^2}{2!} + \ldots + (-1)^k \frac{x^{2k}}{(2k)!} + \ldots \qquad (8.107)$$

we can define the elementary functions e^z, $\cos z$ and $\sin z$ of a complex variable z which coincide, respectively, with e^x, $\cos x$ and $\sin x$ for $z = x$. We remind the reader that series (8.105)-(8.107) are convergent for all real values of x. Hence, by the above lemma, they converge for all complex values of z if z is substituted for x into

---

* When applying D'Alembert's test and Cauchy's root test to a series whose terms are not positive real numbers one should take the moduli of the terms.

** This lemma is a generalization of the one proved for power series in a real argument x (see the proof of Theorem 8.5). The basic idea of the proof remains the same for the complex argument $z = x + iy$.

the series. Therefore, putting

$$e^{z}=1+z+\frac{z^2}{2!}+\ldots+\frac{z^n}{n!}+\ldots \tag{8.108}$$

$$\sin z=z-\frac{z^3}{3!}+\ldots+(-1)^k\frac{z^{2k+1}}{(2k+1)!}+\ldots \tag{8.109}$$

and

$$\cos z=1-\frac{z^2}{2!}+\ldots+(-1)^k\frac{z^{2k}}{(2k)!}+\ldots \tag{8.110}$$

we obtain the corresponding functions of the complex variable $z=x+iy$ defined for all z. By the above lemma, the series thus obtained are absolutely convergent. Therefore it is permissible to perform the arithmetic operations of addition, subtraction and multiplication on them. This enables us to establish the following identities for the functions e^z, $\cos z$ and $\sin z$ of the complex variable z:

$$e^{z_1}\cdot e^{z_2}=e^{z_1+z_2}\text{*} \tag{8.111}$$

and

$$\cos^2 z+\sin^2 z=1 \tag{8.112}$$

Substituting iz for z into (8.108), grouping separately the terms containing and not containing i (after i^2 has been replaced by -1 in all the terms) and making use of formulas (8.109) and (8.110) we arrive at Euler's formula

$$e^{iz}=\cos z+i\sin z \tag{8.113}$$

valid for every complex z. Indeed, we have

$$e^{iz}=1+iz+\frac{i^2z^2}{2!}+\frac{i^3z^3}{3!}+\frac{i^4z^4}{4!}+\ldots=1+iz+\frac{i^2z^2}{2!}+\frac{i\cdot i^2z^3}{3!}+$$
$$+\frac{(i^2)^2z^4}{4!}+\ldots=\left(1-\frac{z^2}{2!}+\frac{z^4}{4!}-\ldots\right)+i\left(z-\frac{z^3}{3!}+\frac{z^5}{5!}-\ldots\right)=$$
$$=\cos z+i\sin z$$

Relations (8.109) and (8.110) implying $\cos(-z)=\cos z$ and $\sin(-z)=-\sin z$, we obtain (by substituting $-z$ for z) the relation

$$e^{-iz}=\cos z-i\sin z \tag{8.114}$$

From relations (8.113) and (8.114) we derive

$$\cos z=\frac{e^{iz}+e^{-iz}}{2}\quad\text{and}\quad\sin z=\frac{e^{iz}-e^{-iz}}{2i} \tag{8.115}$$

The last formulas are also spoken of as *Euler's formulas*.

---

* Equality (8.111) results from multiplying the power series for e^{z_1} and e^{z_2} according to the rules that have been proved for a real argument but remain applicable in the case of a complex argument as well.

It follows from formulas (8.115) that the functions $\cos z$ and $\sin z$ can assume arbitrarily large values in the complex plane. For instance, putting $z = -in$ where n is a natural number we obtain

$$\cos(-in) = \frac{e^n + e^{-n}}{2} \to +\infty \quad \text{for } n \to +\infty$$

But nevertheless formula (8.112) remains true in this case.

Power series in a complex argument make it possible to introduce many other functions defined in domains lying in the complex plane such as $\ln(1+z)$, $\arctan z$ and the like.

The theory of functions of a complex variable is one of the most important divisions of modern mathematics and is widely applied in mathematical physics.

Functions of a complex variable enable us to elucidate many paradoxes of the theory of real functions. For instance, the left-hand side of the equality

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \ldots \tag{$*$}$$

is a bounded continuous function defined over the whole real axis whereas the series on the right hand side diverges for $|x| \geqslant 1$. But if we consider the equality ($*$) for the complex values of x we see that the left-hand side of the equality turns into infinity for $x = i$ and hence the point $x = i$ must lie on the circumference of the circle of convergence with centre at the origin. Indeed, if the point $x = i$ were placed inside the circle of convergence the function $\frac{1}{1+x^2}$ would be continuous at the point, which is impossible since it approaches infinity at this point.

In § 4 we considered the function

$$\varphi(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

which has derivatives of all the orders with respect to x at the origin but cannot be expanded into a series in nonnegative integral powers of x. This fact can be explained if we investigate the function

$$\varphi(z) = e^{-\frac{1}{z^2}} \quad (z \neq 0)$$

under the assumption that z can take on all the possible complex values. Substituting $z = iy$ we obtain $e^{-\frac{1}{z^2}} = e^{\frac{1}{y^2}} \to +\infty$ for $y \to 0$ whereas $e^{-\frac{1}{x^2}} \to 0$ for $x \to 0$. Consequently, the definition of this function cannot be extended to the origin so that the function should become continuous. If a power series convergent to $\varphi(x)$

on an interval $-R < x < R$ existed the substitution of z for x would result in a power series converging to $\varphi(z)$ in the circle $|z| < R$ and the function $\varphi(z)$ would be continuous and even differentiable with respect to z at the point $z = 0$, which contradicts the fact that it has a discontinuity at that point. An exhaustive investigation of this example can only be performed by applying the theory of functions of a complex variable.

§ 6. CONVERGENCE IN THE MEAN

In some divisions of mathematics and its applications a measure of closeness of a function $f(x)$ to a function $g(x)$ is interpreted in an "integral" sense. In such an interpretation the functions may be regarded as being close to each other despite the fact that the

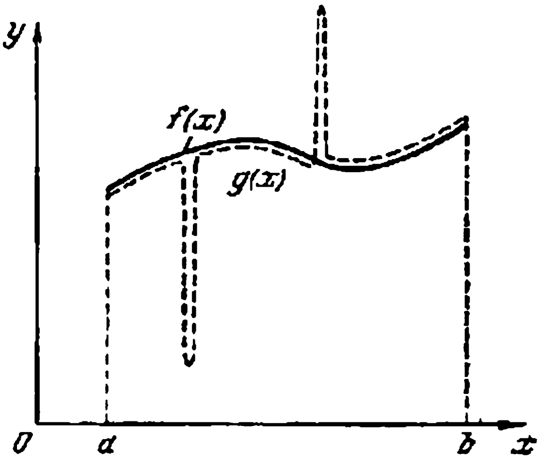


Fig. 8.5

absolute value of their difference $f(x) - g(x)$ takes on large values at separate points. Usually the so-called *mean square deviation* is taken as such a measure, which leads to the notion of *convergence in the mean.*

1. Mean Square Deviation and Convergence in the Mean.

Definition 1. *Given two functions $f(x)$ and $g(x)$, the nonnegative quantity*

$$\rho^2(f, g) = \int_a^b |f(x) - g(x)|^2 \, dx^* \qquad (8.116)$$

is called the ***mean square deviation*** *of the function $f(x)$ from the function $g(x)$ on the interval $[a, b]$.*

---

* In § 6 we suppose that all the functions under consideration are integrable (in the ordinary sense) although most of the assertions and notions introduced here also remain valid for the square integrable (summable) functions (i.e. the ones for which the integrals of the squares of their moduli exist as proper or improper integrals).

We apparently have

$$\rho^2(f, g) = \rho^2(g, f)$$

The graphs of two functions $f(x)$ and $g(x)$ which are close to each other in the sense of their mean square deviation may considerably diverge at separate points (see Fig. 8.5).

Definition 2. A functional sequence

$$f_1(x),\ f_2(x),\ \ldots,\ f_n(x),\ \ldots \tag{8.117}$$

*is said to be **convergent in the mean (in the mean square)** to a function $f(x)$ on an interval $[a, b]$ if*

$$\rho^2(f_n, f) = \int_a^b |f_n(x) - f(x)|^2\,dx \to 0 \quad \text{for } n \to +\infty \tag{8.118}$$

This kind of convergence will be designated by the symbolic relation

$$\lim_{n\to+\infty} f_n(x) \doteq f(x) \quad \text{on} \quad [a, b] \tag{8.119}$$

Definition 3. A functional series

$$\sum_{k=1}^{+\infty} u_k(x) \tag{8.120}$$

*is said to **converge in the mean (square)** to $S(x)$ on an interval $[a, b]$ if the sequence of its partial sums*

$$S_n(x) = \sum_{k=1}^{n} u_k(x), \quad n = 1, 2, \ldots \tag{8.121}$$

is convergent in the mean to $S(x)$ on the interval $[a, b]$, that is

$$\rho^2(S(x), S_n(x)) = \int_a^b \left[S(x) - \sum_{k=1}^{n} u_k(x) \right]^2 dx \to 0 \text{ as } n \to +\infty \tag{8.122}$$

In such a case we shall write

$$S(x) \doteq \sum_{k=1}^{+\infty} u_k(x) \text{ on } [a, b] \tag{8.123}$$

2. **Cauchy-Bunyakovsky[*] Inequality**. *If two functions $f(x)$ and $g(x)$ fulfill the above requirements on $[a, b]$ they satisfy the Cau-*

---

[*] Bunyakovsky, Victor Yakovlevich (1804-1889), a Russian mathematician. This inequality is also spoken of as the *Cauchy-Schwarz inequality* although Bunyakovsky called attention to it earlier than Schwarz (in 1859).

chy-Bunyakovsky-Schwarz inequality

$$\left|\int_a^b f(x)\,g(x)\,dx\right| \leqslant \sqrt{\int_a^b f^2(x)\,dx}\cdot\sqrt{\int_a^b g^2(x)\,dx} \quad (8.124)$$

Proof. Let us put

$$A=\int_a^b f^2(x)\,dx,\quad B=\int_a^b f(x)\,g(x)\,dx \quad \text{and} \quad C=\int_a^b g^2(x)\,dx \quad (8.125)$$

and consider the two cases which are possible here: (1) $A = C = 0$ and (2) at least one of the numbers A and C is different from zero.

(1) If $A = C = 0$, i.e.

$$\int_a^b f^2(x)\,dx=\int_a^b g^2(x)\,dx=0$$

it follows, on the basis of the obvious inequality

$$|f(x)\,g(x)| \leqslant \frac{1}{2}[f^2(x)+g^2(x)]$$

that

$$\int_a^b |f(x)\,g(x)|\,dx \leqslant \frac{1}{2}\left[\int_a^b f^2(x)\,dx+\int_a^b g^2(x)\,dx\right] = 0$$

But we have

$$\left|\int_a^b f(x)\,g(x)\,dx\right| \leqslant \int_a^b |f(x)\,g(x)|\,dx$$

and consequently $B=\int_a^b f(x)\,g(x)\,dx=0$. Therefore inequality (8.124) is fulfilled in this case because its right-hand and left-hand sides are equal to zero.

(2) For definiteness, let $A > 0$. Then we apply the following technique. Taking a scalar parameter λ we can write the relation

$$[\lambda f(x)+g(x)]^2 \geqslant 0$$

which holds for all real values of the parameter. Thus, we have

$$\lambda^2 f^2(x)+2\lambda f(x)\,g(x)+g^2(x) \geqslant 0$$

Integrating the last inequality with respect to x from a to b and making use of notation (8.125) we conclude that the inequality

$$A\lambda^2+2B\lambda+C \geqslant 0 \quad (8.126)$$

(in which $A > 0$) holds for all real values of λ. Consequently the quadratic trinomial $A\lambda^2 + 2B\lambda + C$ cannot have two distinct real roots $\lambda_1 < \lambda_2$ because if otherwise it could be represented in the form $A(\lambda - \lambda_1)(\lambda - \lambda_2)$ and hence would assume negative values for λ satisfying the condition $\lambda_1 < \lambda < \lambda_2$, which contradicts inequality (8.126). But, as is well known, for a quadratic trinomial to have no distinct real roots it is necessary and sufficient that the discriminant of the trinomial be nonpositive:

$$B^2 - AC \leqslant 0 \tag{8.127}$$

Transposing the product AC to the right-hand side of the inequality and extracting the square root of either side we obtain $|B| \leqslant \leqslant \sqrt{|A|}\cdot\sqrt{|C|}$. Taking into account notation (8.125) we thus arrive at the Cauchy-Bunyakovsky inequality.

3. Integration of Sequences and Series Convergent in the Mean.

Theorem 8.14$_1$. *If a functional sequence $f_1(x), f_2(x), \ldots, f_n(x), \ldots$ converges in the mean to a function $f(x)$ on an interval $[a, b]$ we have, for any x_0 and x belonging to $[a, b]$, the relation*

$$\lim_{n\to+\infty} \int_{x_0}^{x} f_n(z)\,dz = \int_{x_0}^{x} f(z)\,dz\,^* \tag{8.128}$$

Moreover, for any $x_0 \in [a, b]$, the sequence $\left\{\int_{x_0}^{x} f_n(x)\,dx\right\}$, $n = 1, 2, \ldots$ (regarded as a sequence of functions dependent on x) uniformly converges to $\int_{x_0}^{x} f(z)\,dz$, i.e.

$$\int_{x_0}^{x} f_n(z)\,dz \rightrightarrows \int_{x_0}^{x} f(z)\,dz \text{ on } [a, b] \tag{8.129}$$

Proof. We have, by the hypothesis, the relation

$$\rho^2(f, f_n) = \int_a^b [f(z) - f_n(z)]^2\,dz \to 0, \quad \text{for } n \to +\infty \tag{8.130}$$

To prove (8.129) (and, consequently, (8.128) as well) we shall estimate the integral $\int_{x_0}^{x} [f(z) - f_n(z)]\,dz$. For this purpose we rep-

* We remind the reader that all the functions (including $f_n(x)$ and $f(x)$), considered in § 6 are supposed to be integrable on $[a, b]$ in the ordinary sense. See also footnote on page 366.

resent the integrand as the product

$$[f(z)-f_n(z)] = 1\cdot[f(z)-f_n(z)]$$

and apply the Cauchy-Bunyakovsky inequality:

$$\left|\int_{x_0}^{x}[f(z)-f_n(z)]\,dz\right| \leqslant \sqrt{\int_{x_0}^{x}1^2\,dz}\cdot\sqrt{\int_{x_0}^{x}|f(z)-f_n(z)|^2\,dz} \leqslant$$

$$\leqslant \sqrt{b-a}\cdot\sqrt{\int_a^b |f(z)-f_n(z)|^2\,dz} =$$

$$= \sqrt{b-a}\,\sqrt{\rho^2(f, f_n)} \qquad (8.131)$$

The right-hand side of inequality (8.131) being independent of x and tending to zero as $n\to+\infty$ (by relation (8.130)), we can write

$$\int_{x_0}^{x}[f(z)=f_n(z)]\,dz \rightrightarrows 0 \text{ on } [a, b]$$

which means that relation (8.129) is fulfilled. Thus, the proof of the theorem has been completed.

If we take the modulus of the difference $|f(x)-f_n(x)|$ instead of $f(x)-f_n(x)$ and put $x_0=a$ and $x=b$ inequality (8.131) is replaced by the relation

$$\int_a^b |f(z)-f_n(z)|\,dz \leqslant \sqrt{b-a}\cdot\sqrt{\rho^2(f, f_n)} \qquad (8.132)$$

The integral on the left-hand side of (8.132) expresses the area lying between the graphs of $f_n(x)$ and $f(x)$ and bounded on the left and on the right by the respective vertical straight lines $x=a$ and $x=b$. Hence, in case $f_n(x)$ converges in the mean to $f(x)$ on the interval $[a, b]$ the area tends to zero since the mean square deviation $\rho^2(f, f_n)$ approaches zero. But at the same time the **maximum deviation** of $f_n(x)$ from $f(x)$ on $[a, b]$, that is the quantity $\sup\limits_{a\leqslant x\leqslant b}|f_n(x)-f(x)|$, may infinitely increase since the values of $f_n(x)$ and $f(x)$ may considerably differ for any n at some separate points. For instance, the sequence $f_n(x)=n^{-\frac{1}{8}}\sqrt{2nx}\,e^{-\frac{1}{2}nx^2}$ converges in the mean to $f(x)\equiv 0$ on the closed interval $0\leqslant x\leqslant 1$ but

$$\max_{0\leqslant x\leqslant 1}|f_n(x)-f(x)| = f_n\left(\frac{1}{\sqrt{2n}}\right) = \sqrt[4]{\frac{2}{e}}\,n^{\frac{1}{8}} \to +\infty \text{ for } n\to+\infty$$

Taking a series convergent in the mean we can apply Theorem 8.14_1 to the sequence of its partial sums and thus obtain the corresponding theorem concerning series:

Theorem 8.14₂. If a functional series $\sum_{k=1}^{+\infty} u_k(x)$ *with integrable terms converges in the mean to an integrable function* $S(x)$ *on an interval* $[a, b]$ *the relation*

$$\int_{x_0}^{x} S(z)\,dz = \int_{x_0}^{x} u_1(z)\,dz + \ldots + \int_{x_0}^{x} u_n(z)\,dz + \ldots \tag{8.133}$$

holds for any x_0 *and* x *belonging to* $[a, b]$ *and series* (8.133) *uniformly converges to its sum on* $[a, b]$.

Proof. By the hypothesis, $S_n(x)$ converges in the mean to $S(x)$ on $[a, b]$ (where $S_n(x) = \sum_{k=1}^{n} u_k(x)$) and therefore, by Theorem 8.14_1, we have

$$\int_{x_0}^{x} S_n(z)\,dz \rightrightarrows \int_{x_0}^{x} S(z)\,dz \text{ on } [a, b] \tag{8.134}$$

But

$$\int_{x_0}^{x} S_n(z)\,dz = \sum_{k=1}^{n} \int_{x_0}^{x} u_k(z)\,dz$$

and consequently

$$\left\{\sum_{k=1}^{n} \int_{x_0}^{x} u_k(z)\,dz\right\} \rightrightarrows \int_{x_0}^{x} S(z)\,dz$$

Hence, the sequence of partial sums of series (8.133) is uniformly convergent to the function $\int_{x_0}^{x} S(z)\,dz$ which is what we set out to prove.

4. Connection Between Convergence in the Mean and Term-by-Term Differentiation of Sequences and Series.

Theorem 8.15₁. If a sequence of continuously differentiable functions $\{f_n(x)\}$ *is convergent in the mean to a function* $f(x)$ *on an interval* $[a, b]$ *and the sequence of the derivatives* $\{f_n'(x)\}$ *is convergent in the mean to a continuous function* $\varphi(x)$ *on* $[a, b]$ *the function* $f(x)$ *is differentiable on* $[a, b]$ *and*

$$f'(x) = \varphi(x) \doteq \lim_{n\to+\infty} f_n'(x) \quad \text{on } [a, b] \tag{8.135}$$

Proof. We have, for x, $x_0 \in [a, b]$, the inequality

$$\left| f_n(x) - f_n(x_0) - \int_{x_0}^{x} \varphi(z)\, dz \right| = \left| \int_{x_0}^{x} [f_n'(z) - \varphi(z)]\, dz \right| \leqslant$$

$$\leqslant \sqrt{\int_{x_0}^{x} 1^2\, dz} \cdot \sqrt{\int_{x_0}^{x} [f_n'(z) - \varphi(z)]^2\, dz} \leqslant$$

$$\leqslant \sqrt{b-a} \cdot \sqrt{\rho^2(f_n', \varphi)} \to 0 \qquad (8.136)$$

as $n \to +\infty$. Consequently, passing to the limit and taking into account that $f_n(x) \to f(x)$ and $f_n(x_0) \to f(x_0)$ as $n \to +\infty$ we obtain

$$f(x) - f(x_0) = \int_{x_0}^{x} \varphi(z)\, dz, \text{ i.e. } f(x) = f(x_0) + \int_{x_0}^{x} \varphi(z)\, dz \qquad (8.137)$$

It follows from equality (8.137) that $f(x)$ is differentiable and that equality (8.135) takes place. The theorem has been proved.

We similarly prove

Theorem 8.15$_2$. *If a functional series*

$$S(x) = \sum_{k=1}^{+\infty} u_k(x) \qquad (8.138)$$

with continuously differentiable terms is convergent on $[a, b]$ *and the series*

$$\sigma(x) \doteq \sum_{k=1}^{+\infty} u_k'(x) \qquad (8.139)$$

converges in the mean to a continuous function $\sigma(x)$ *the sum* $S(x)$ *of series* (8.138) *is differentiable on* $[a, b]$ *and*

$$S'(x) = \sigma(x) \doteq \sum_{k=1}^{+\infty} u_k'(x) \qquad (8.140)$$

The proof of the theorem is left to the reader.

5. Connection Between Convergence in the Mean and Other Types of Convergence. If a sequence of functions $f_1(x)$, $f_2(x)$, ..., $f_n(x)$, ... converges (in the ordinary sense) at each point of an interval $[a, b]$ this does not imply, in the general case, its convergence in the mean. For example, we have $f_n(x) = \sqrt{2nx}\, e^{-\frac{1}{2}nx^2} \to f(x) = 0$ on the interval $[0, 1]$ as $n \to +\infty$ whereas

$$\int_0^1 [f_n(x) - f(x)]^2\, dx = \int_0^1 2nxe^{-nx^2}\, dx = (1 - e^{-n}) \to 1 \quad \text{for } n \to +\infty$$

But if a sequence $\{f_n(x)\}$ uniformly converges to a function $f($ on $[a, b]$ it necessarily converges in the mean to the same functio In fact, if for any $\varepsilon > 0$ there is $N(\varepsilon)$ such that the inequalit $|f_n(x) - f(x)| < \sqrt{\frac{\varepsilon}{b-a}}$ is fulfilled for all $n > N(\varepsilon)$ an all $x \in [a, b]$ we can square both sides of the inequality and integrat the result, which yields the relation

$$\rho^2(f, f_n) = \int_a^b |f_n(x) - f(x)|^2 \, dx < \frac{\varepsilon}{b-a}(b-a) = \varepsilon$$

valid for all $n > N(\varepsilon)$. Hence $\rho^2(f, f_n) \to 0$, as $n \to +\infty$, whic implies that $f_n(x)$ is convergent in the mean to $f(x)$ on $[a, b]$.

Uniform convergence does not follow from convergence in th mean; moreover, a sequence convergent in the mean on an interva

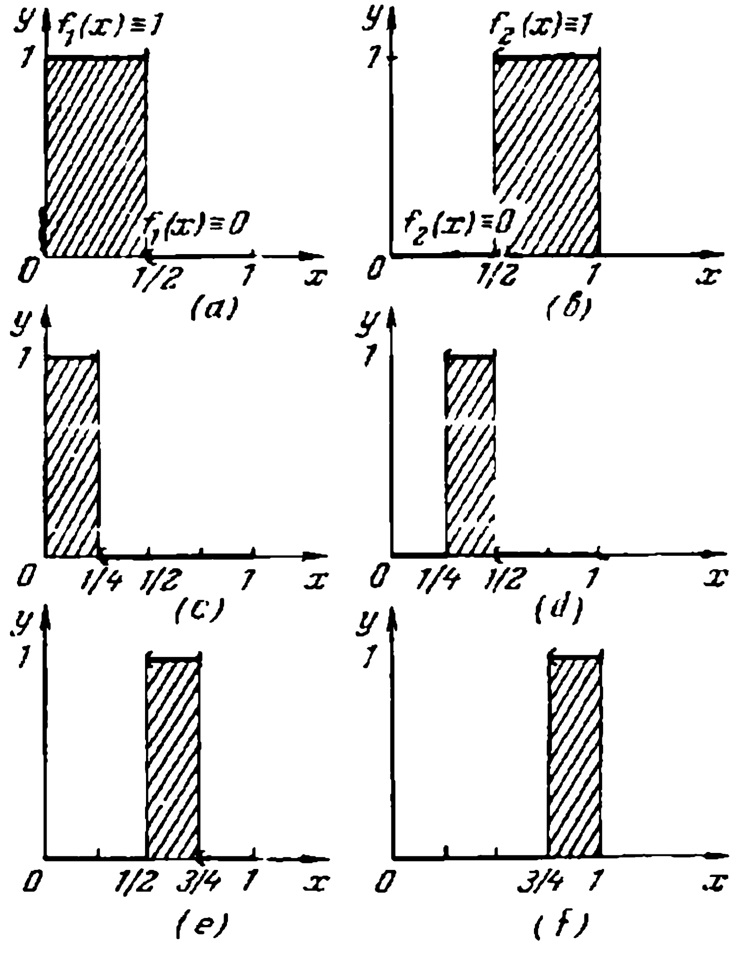


Fig. 8.6

may even diverge at each point of the interval. As an example, le us construct a sequence of functions $f_1(x), f_2(x), \ldots, f_n(x), \ldots$ convergent in the mean to $f(x) = 0$ on the interval $[0, 1]$ which diver ges at every point of the interval. To this end, we first divide th interval $[0, 1]$ into 2 equal parts and define the functions $f_1(x$ and $f_2(x)$ as is shown in Fig. 8.6*a* and *b*. The graphs of the function are given in heavy lines and the small arcs indicate that the corres

ponding points do not belong to the adjoining pieces of the graphs.

Next we divide the interval [0, 1] into 2^2 equal parts and define the functions $f_3(x)$, $f_4(x)$, $f_5(x)$ and $f_6(x)$ in a manner shown in Fig. 8.6*c*, *d*, *e* and *f*. Continuing infinitely in this way we obtain a functional sequence $f_1(x), f_2(x), \ldots, f_n(x), \ldots$ whose every member $f_n(x)$ takes on only two values, namely 0 and 1, on the interval [0, 1]. Therefore $f_n^2(x) = f_n(x)$ for all $n = 1, 2, 3, \ldots$ on the interval [0, 1].

Now let us prove that $f_n(x)$ converges in the mean to $f(x) \equiv 0$ on [0, 1]. We have

$$\rho^2(f, f_n) = \int_0^1 |f_n(x) - f(x)|^2\,dx = \int_0^1 f_n^2(x)\,dx = \int_0^1 f_n(x)\,dx \to 0$$

for $n \to +\infty$. Indeed, the integral $\int_0^1 f_n(x)\,dx$ is equal to the area shaded in Fig. 8.6, and when $n \to +\infty$ this area obviously tends to zero. Convergence in the mean has thus been proved.

On the other hand, the functional sequence $f_1(x), f_2(x), \ldots, f_n(x), \ldots$ thus constructed diverges at each point of the interval [0, 1] because for every fixed point x belonging to the interval and for an arbitrarily large $N > 0$ there are functions $f_{n'}(x)$ and $f_{n''}(x)$ with $n', n'' > N$ such that the former assumes the value $f_{n'}(x) = 0$ and the latter the value $f_{n''}(x) = 1$ at the point x.

APPENDIX I TO CHAPTER 8

CRITERION FOR COMPACTNESS OF A FAMILY OF FUNCTIONS

The notion of *compactness* of a family of functions appears in mathematical physics in connection with theorems on existence of solutions of differential and integral equations and on convergence of various approximating processes for evaluating solutions of such equations. But the definition of this notion does not involve any concepts connected with the theory of differential or integral equations and is closely related to the questions discussed in the present chapter.

Definition 1. A family of functions $\{f(x)\}$ defined on a set X of points x is said to be **compact** *(in the sense of uniform convergence) if for every infinite sequence of functions $f_1(x), f_2(x), \ldots, f_n(x), \ldots$ belonging to the family there exists a subsequence*

$f_{n_1}(x), f_{n_2}(x), \ldots, f_{n_k}(x), \ldots$ which is uniformly convergent on the set X. *

*Definition 2. A family of functions $\{f(x)\}$ defined on a set X is said to be **uniformly bounded** on this set if there is a constant C, $0 < C < +\infty$ such that the inequality $|f(x)| \leqslant C$ is fulfilled for all $x \in X$ and all functions $f(x)$ belonging to the family.*

*Definition 3. We say that a family of functions $\{f(x)\}$ defined on a set X is **equicontinuous** on the set if, given an arbitrary $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that for every function $f(x) \in \{f(x)\}$ and any $x', x'' \in X$ satisfying the condition $|x' - x''| < \delta(\varepsilon)$ we have the inequality*

$$|f(x') - f(x'')| < \varepsilon \tag{1}$$

where $\delta(\varepsilon)$ is independent of the choice of x' and x'' and is determined solely by the value of ε.

*Theorem 1 (Arzelà's Criterion for Compactness**). If a family of functions $\{f(x)\}$ defined on a closed interval $a \leqslant x \leqslant b$ is uniformly bounded and equicontinuous it is compact (in the sense of uniform convergence).*

Proof. Take an arbitrary countable set M of points $x_1, x_2, \ldots, x_n, \ldots$ (belonging to the interval $a \leqslant x \leqslant b$) which is everywhere dense in $[a, b]$.*** For instance, we can take, as $x_1, x_2, \ldots, x_n, \ldots$, the set of all rational points of the interval or the set of all the points appearing in the process of successive division of the interval into $2, 4, 8, \ldots, 2^n, \ldots$ equal parts. Let us choose a sequence of functions $\{f_n(x)\}$ belonging to the family. The family being uniformly bounded, we have a relation $|f_n(x)| \leqslant C = \text{const} < +\infty$ for all $n = 1, 2, 3, \ldots$ and all $x \in M$. In particular, the numerical sequence $f_n(x_1)$, $n = 1, 2, \ldots$, is bounded and hence, by the Bolzano-Weierstrass theorem, there is a convergent subsequence of this numerical sequence whose members can be numbered as $f_{11}(x_1), f_{12}(x_1), \ldots, f_{1n}(x_1), \ldots$. Thus, we have

---

* Uniform convergence on an arbitrary set is defined in a manner completely analogous to that of defining the notion of uniform convergence on an interval. In functional analysis the concept of compactness is also introduced. in a similar fashion, for convergence in the mean and for other kinds of convergence but we shall not dwell on these questions here.

** Also known as the Ascoli-Arzelà theorem. Ascoli Giulio (1843-1896) and Arzelà, Cesare (1847-1912). Italian mathematicians.

*** A subset B of a point set A is said to be (everywhere) dense in A if the closure of B coincides with A (compare with footnote on page 29).

taken, from the sequence $\{f_n(x)\}$, the functional subsequence

$$f_{11}(x),\ f_{12}(x),\ \ldots,\ f_{1n}(x),\ \ldots \tag{2_1}$$

convergent at the point $x_1 \in M$.

We can now similarly construct a subsequence

$$f_{21}(x),\ f_{22}(x),\ \ldots,\ f_{2n}(x),\ \ldots \tag{2_2}$$

of sequence (2_1) convergent at the point $x_2 \in M$. Subsequence (2_2) is also convergent at the point $x_1 \in M$ since its members belong to the sequence (2_1), the latter being convergent at the point x_1. Thus, subsequence (2_2) converges at the points x_1 and x_2 belonging to the set M.

Continuing unlimitedly this process of selecting subsequences we arrive at an infinite table of the form

$$\begin{array}{l} f_{11}(x),\ f_{12}(x),\ \ldots,\ f_{1n}(x),\ \ldots \\ f_{21}(x),\ f_{22}(x),\ \ldots,\ f_{2n}(x),\ \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_{n1}(x),\ f_{n2}(x),\ \ldots,\ f_{nn}(x),\ \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array} \tag{3}$$

whose nth row ($n = 1, 2, 3, \ldots$) is a subsequence (of the original sequence $\{f_n(x)\}$) converging at the points $x_1, x_2, \ldots, x_n$. Therefore the "diagonal" subsequence

$$f_{11}(x),\ f_{22}(x),\ \ldots,\ f_{nn}(x),\ \ldots \tag{4}$$

converges at each point $x_n \in M$.[1]

Let us prove that this sequence is uniformly convergent on $[a, b]$. To this end, it is sufficient to show that sequence (4) satisfies the condition of Cauchy's test for uniform convergence of a sequence. Take an arbitrary $\varepsilon > 0$. Let us choose $\delta(\varepsilon) > 0$ such that the condition

$|f(x') - f(x'')| < \varepsilon$ for $|x' - x''| < \delta(\varepsilon)$ and $x', x'' \in [a, b]$ holds, which is always possible since the family is equicontinuous. Next take a finite subset $\widetilde{M}\,[x_1, x_2, \ldots, x_p]$ of the dense set M whose points break up the interval $[a, b]$ into subintervals of lengths less than $\delta(\varepsilon)$. Then, for each $x \in [a, b]$, there exists a value $x_i \in \widetilde{M}$ such that $|x - x_i| < \delta(\varepsilon)$. Furthermore, for the given $\varepsilon > 0$, let us take $N(\varepsilon)$ independent of x_i, $i = 1, 2, \ldots, p$, such that

$$|f_{mm}(x_i) - f_{nn}(x_i)| < \varepsilon \quad \text{for all } m, n > N(\varepsilon) \tag{5}$$

and all $i = 1, 2, \ldots, p$. Then we have the inequality $|f_{mm}(x) - f_{nn}(x)| < 3\varepsilon$ for all $m, n > N(\varepsilon)$ and all $x \in [a, b]$. Indeed, suppose $m, n > N(\varepsilon)$ and $x \in [a, b]$. Then there is a value $x_i \in \widetilde{M}$

such that $|x - x_i| < \delta(\varepsilon)$. The family being equicontinuous, it follows that $|f_{mm}(x) - f_{mm}(x_i)| < \varepsilon$ and $|f_{nn}(x) - f_{nn}(x_i)| < \varepsilon$. Consequently, by virtue of (5), the inequality

$$|f_{mm}(x) - f_{nn}(x)| \leqslant |f_{mm}(x) - f_{mm}(x_i)| + |f_{mm}(x_i) - f_{nn}(x_i)| + \\ + |f_{nn}(x_i) - f_{nn}(x)| < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon$$

is fulfilled. Hence, sequence (4) satisfies the condition of Cauchy's test for uniform convergence and thus the theorem has been proved.

Arzelà's theorem has many applications in various divisions of mathematics. In particular, it is used for proving existence theorems of the theory of differential equations.

Let us prove another theorem widely applied in the theory of integral equations.

Theorem 2. If a functional sequence $\{f_n(x)\}$ *is equicontinuous on* $[a, b]$ *and satisfies the condition*

$$\rho^2(f_n, f_m) = \int_a^b [f_m(x) - f_n(x)]^2 \, dx \to 0 \quad \text{for } n, m \to +\infty \tag{6}$$

it is uniformly convergent on the interval $[a, b]$ *to a continuous function* $f(x)$.

Proof. We shall show that such a sequence satisfies the condition of Cauchy's test for uniform convergence, that is

$$\varphi_{n,m}(x) = |f_n(x) - f_m(x)| \rightrightarrows 0 \quad \text{on } [a, b] \quad \text{as } n, m \to +\infty$$

Assume the contrary. Then there is $\varepsilon_0 > 0$ such that for an arbitrarily large k there exist $n_k > k$ and $x_k \in [a, b]$ for which

$$|\varphi_{k,n_k}(x_k)| \geqslant 4\varepsilon_0 \tag{7}$$

The sequence being equicontinuous, there is $\delta = \delta(\varepsilon_0)$ such that

$$|f_n(x) - f_n(x_k)| < \varepsilon_0 \quad \text{for } |x - x_k| < \delta(\varepsilon_0) \tag{8}$$

Consequently, for $n_k > k > N(\varepsilon_0)$ and $|x - x_k| < \delta(\varepsilon_0)$ we have

$$|\varphi_{k,n_k}(x) - \varphi_{k,n_k}(x_k)| \leqslant |f_k(x) - f_k(x_k)| + \\ + |f_{n_k}(x) - f_{n_k}(x_k)| < 2\varepsilon_0 \tag{9}$$

Then, by virtue of (7) and (9), we can write

$$|\varphi_{k,n_k}(x)| \geqslant |\varphi_{k,n_k}(x_k)| - |\varphi_{k,n_k}(x) - \varphi_{k,n_k}(x_k)| \geqslant \\ \geqslant 4\varepsilon_0 - 2\varepsilon_0 = 2\varepsilon_0 \tag{10}$$

If we take $\delta(\varepsilon_0) < b - a$ then at least half of the interval $x_k - \delta < x < x_k + \delta$ belongs to the interval $[a, b]$. Hence, by (10)

the inequality

$$\rho^2(f_{n_k}, f_k) = \int_a^b \varphi^2_{k,\,n_k}(x)\,dx > 4\varepsilon_0^2\cdot\frac{\delta}{2} = 2\varepsilon_0^2\delta = \text{const} > 0 \qquad (11)$$

is fulfilled for $n_k > k > N(\varepsilon_0)$, which contradicts condition (6) since the number k can be chosen arbitrarily large. Therefore $\varphi_{m,\,n}(x) = [f_m(x) - f_n(x)] \rightrightarrows 0$ on $[a, b]$ as $n, m \to +\infty$. Consequenly, by virtue of Cauchy's test for uniform convergence of a sequence, the sequence $\{f_n(x)\}$ is uniformly convergent on $[a, b]$ to a function $f(x)$ which is continuous as a limit of a uniformly converging sequence of continuous functions. The theorem has been proved.

APPENDIX 2 TO CHAPTER 8

WEAK CONVERGENCE AND DELTA FUNCTION

Besides uniform convergence and convergence in the mean, in mathematics and mathematical physics the so-called *weak convergence* also plays an important role.

Definition 1. *A sequence of functions*

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots \qquad (1)$$

defined and integrable on an interval (a, b) *is said to be a* ***weakly fundamental sequence*** *on the interval* (a, b) *if for every bounded and continuous function* $f(x)$ *there exists a finite limit*

$$\lim_{n\to+\infty}\int_a^b f(x)\,\varphi_n(x)\,dx \qquad (2)$$

Definition 2. *A function* $\varphi(x)$ *is called a* ***weak limit*** *of functional sequence* (1) *(or the sequence is said to be* ***weakly convergent to*** $\varphi(x)$*) on an interval* (a, b) *if the relation*

$$\lim_{n\to+\infty}\int_a^b f(x)\,\varphi_n(x)\,dx = \int_a^b f(x)\,\varphi(x)\,dx \qquad (3)$$

holds for every bounded continuous function $f(x)$ *defined on* (a, b).*

If sequence (1) is uniformly convergent or convergent in the mean to an integrable function $\varphi(x)$ equality (3) is fulfilled for every

* More precisely, $\varphi(x)$ is a weak limit of sequence (1) relative to the class of functions $f(x)$ continuous on (a, b). The notion of weak convergence can also be defined for other classes of functions.

bounded continuous function $f(x)$ because this is implied by the theorems on passing to limit under the sign of integration. Thus, if a sequence $\{\varphi_n(x)\}$ converges uniformly or converges in the mean to a limit function $\varphi(x)$ it is always weakly convergent to $\varphi(x)$.

If we have

$$\lim_{n\to+\infty}\int_a^b f(x)\,\varphi_n(x)\,dx=0 \tag{4}$$

for any continuous bounded function $f(x)$ defined on (a, b) the sequence $\{\varphi_n(x)\}$ is weakly convergent to zero on (a, b) since equality (4) can be rewritten in the form

$$\lim_{n\to+\infty}\int_a^b f(x)\,\varphi_n(x)\,dx=\int_a^b f(x)\cdot 0\cdot dx \tag{4'}$$

Applying Cauchy's test to the number sequence $\left\{\int_a^b f(x)\,\varphi_n(x)\,dx\right\}$ we arrive at the following test for a functional sequence to be weakly fundamental:

Cauchy's Test (for a Sequence to Be Weakly Fundamental). *Sequence* (1) *is weakly fundamental on* (a, b) *if and only if for every* $\varepsilon>0$ *and every continuous function* $f(x)$ *there exists a number* $N(\varepsilon, f)$ *such that*

$$\left|\int_a^b f(x)\,[\varphi_n(x)-\varphi_m(x)]\,dx\right|<\varepsilon \tag{5}$$

for all $n, m>N(\varepsilon, f)$.

We remind the reader that a fundamental (Cauchy) sequence of rational numbers may not converge to a rational number and this leads to the irrational numbers which enlarge the class of rational numbers so that every fundamental number sequence has a limit. Similarly, there are weakly fundamental sequences of integrable functions which do not weakly converge to an integrable function and therefore it is advisable to enlarge the class of functions under consideration in an appropriate manner. This leads to the so-called generalized functions.

For instance, let us consider a sequence of functions $\{\delta_n(x_0, x)\}$ determined by the relations

$$\delta_n(x_0, x)=\begin{cases} n \text{ for } x_0-\frac{1}{2n}<x<x_0+\frac{1}{2n}\\ 0 \text{ for } -\infty<x<x_0-\frac{1}{2n} \text{ and } x_0+\frac{1}{2n}<x<-\infty\end{cases} \tag{6}$$

The sequence is weakly fundamental on every interval (a, b).

Indeed, let a function $f(x)$ be continuous on (a, b) and $x_0 \in (a, b)$. Then, beginning with a sufficiently large n, the interval $\left(x_0 - \frac{1}{2n},\ x_0 + \frac{1}{2n}\right)$ is contained within (a, b) and hence, applying the mean value theorem to the integral $\int\limits_a^b f(x)\,\delta_n(x_0, x)\,dx$, we obtain

$$\int\limits_a^b f(x)\,\delta_n(x_0, x)\,dx = n \int\limits_{x_0-\frac{1}{2n}}^{x_0+\frac{1}{2n}} f(x)\,dx = f(\xi),$$
$$x_0 - \frac{1}{2n} \leqslant \xi \leqslant x_0 + \frac{1}{2n} \tag{6'}$$

Passing to the limit in the last equality, as $n \to +\infty$, we derive, by the continuity of $f(x)$, the relation

$$\lim_{n\to+\infty} \int\limits_a^b f(x)\,\delta_n(x_0, x)\,dx = f(x_0) \tag{7}$$

In case x_0 lies outside the closed interval $[a, b]$ we find

$$\lim_{n\to+\infty} \int\limits_a^b f(x)\,\delta_n(x_0, x)\,dx = 0 \tag{8}$$

But of course there is no ordinary integrable function which is a weak limit of the sequence $\{\delta_n(x_0, x)\}$. Therefore, to define the weak limit of this sequence we introduce a *generalized function* $\delta(x_0, x)$ called **δ-function** (the **Dirac* delta functional** or **distribution**).

Making use of the definition of a weak limit (see equality (3)) and taking into account relations (7) and (8) we can write, for every continuous function $f(x)$ on (a, b), the relation

$$\int\limits_a^b f(x)\,\delta(x_0, x)\,dx = \begin{cases} f(x_0) & \text{for } x_0 \in (a, b) \\ 0 & \text{for } x_0 \notin [a, b] \end{cases} \tag{9}$$

Formal relation (9) is sometimes taken as a definition of the delta function.

Note. Instead of a bounded continuous function $f(x)$ defined on (a, b) which enters into relations (7) and (8) we can obviously take an arbitrary bounded piecewise continuous function $f(x)$

---

* Dirac, Paul Adrien Maurice (born in 1902), an English physicist.

if its value at each point of discontinuity x_0 belonging to (a, b) is defined as $f(x_0) = \frac{f(x_0-0)+f(x_0+0)}{2}$. It is therefore natural to extend relation (9) to such functions.*

We know that an irrational number can be determined by means of various equivalent fundamental number sequences of rational numbers. Similarly, a generalized function can be regarded as a limit of different *equivalent* weakly fundamental sequences. To define equivalence we introduce the following

*Definition 3. Two functional sequences $\{\varphi_n(x)\}$ and $\{\psi_n(x)\}$ weakly fundamental on (a, b) are called **equivalent** if the relation*

$$\lim_{n\to+\infty}\int_a^b f(x)\,[\varphi_n(x)-\psi_n(x)]\,dx = 0$$

is fulfilled for every bounded continuous function $f(x)$ defined on (a, b).

In practical applications of the delta function we often take various weakly fundamental sequences (convergent to $\delta(x_0, x)$ in the sense of relation (3)) equivalent to the sequence $\{\delta_n(x_0, x)\}$ used in this section (see Appendix 5 to Chapter 11). A feature of relation (9) which formally defines the delta function is that when we multiply an arbitrary continuous function $f(x)$ by $\delta(x_0, x)$ and integrate the product over an interval in which $f(x)$ is defined and which contains x_0 we "isolate" the value $f(x_0)$ of the function $f(x)$ assumed at the point x_0.

*If to each function belonging to a certain class of functions defined on (a, b) there corresponds a number we say that we have a **functional** defined on that class of functions.*

---

* If x_0 is a point of discontinuity of a piecewise continuous function $f(x)$ and $f(x_0) = \frac{f(x_0-0)+f(x_0+0)}{2}$ then to prove equality (7) for $x_0 \in (a, b)$ it is necessary to break up the integral $\int\limits_{x_0-\frac{1}{2n}}^{x_0+\frac{1}{2n}} f(x)\,dx$ in relation (6′) into the two integrals $\int\limits_{x_0-\frac{1}{2n}}^{x_0} f(x)\,dx$ and $\int\limits_{x_0}^{x_0+\frac{1}{2n}} f(x)\,dx$ and separately apply the mean value theorem to either integral.

For example, let $\{f(x)\}$ be the class of all functions integrable on (a, b). Then the integral $\int_a^b f(x)\,dx$ is a functional defined on the class.

Integral (9) also determines, for every fixed $x_0 \in (a, b)$, a functional defined on the class of all functions continuous on (a, b). Indeed, relation (9) associates, with every continuous function $f(x)$ defined on (a, b), a uniquely specified number, namely the value $f(x_0)$ of this function at the point $x_0 \in (a, b)$.

δ-function is sometimes identified with the functional determined by relation (9). This is one of the possible interpretations of generalized functions which can serve as a basis for the theory of generalized functions.

9 Improper Integrals

The definition of the integral as the limit of the integral sums

$$\int_a^b f(x)\,dx = \lim_{\max \Delta x_k \to 0} \sum_{k=0}^{n} f(\xi_k)\,\Delta x_k \qquad (9.1)$$

does not embrace those cases when the integrand is an unbounded function or the interval of integration is infinite. In mathematics and mathematical physics, however, we often use integrals of unbounded functions or integrals with unbounded domains of integration. Such integrals are called *improper*. To define them it is not sufficient to apply a passage to the limit of type (9.1) but it is necessary to use an additional passage to the limit involving the domain of integration. The original domain of integration where definition (9.1) does not apply is replaced by a subdomain where it holds. Then this subdomain is made to extend so that it tends to coincide with the original domain. The limit of the integral taken over this subdomain is called the improper integral over the original domain. This is the general idea the definition of the improper integral is based on. The strict definition will be given below.

§ 1. INTEGRALS WITH INFINITE LIMITS OF INTEGRATION

1. Definitions. Examples. Let a function $f(x)$ be defined on an interval $a \leqslant x < +\infty$ and let the integral $\int_a^B f(x)\,dx$ (determined by relation (9.1)) exist for every $B > a$.

Definition 1. *The improper integral* $\int_a^{+\infty} f(x)\,dx$ *is understood as the limit*

$$\int_a^{+\infty} f(x)\,dx = \lim_{B \to +\infty} \int_a^B f(x)\,dx \qquad (9.2)$$

If this limit exists and is finite the impoprer integral $\int\limits_a^{+\infty} f(x)\,dx$ *is said to be* ***convergent****. If otherwise it is said to be* ***divergent****.*

Note. Let $a_1 > a$. Then the equality

$$\int\limits_a^B f(x)\,dx = \int\limits_a^{a_1} f(x)\,dx + \int\limits_{a_1}^B f(x)\,dx$$

implies that the integrals $\int\limits_a^{+\infty} f(x)\,dx$ and $\int\limits_{a_1}^{+\infty} f(x)\,dx$ are simultaneously convergent or divergent. Thus, when testing the integral $\int\limits_a^{+\infty} f(x)\,dx$ for convergence, we can replace it by the integral $\int\limits_{a_1}^{+\infty} f(x)\,dx$ for any $a_1 > a$ provided that the function $f(x)$ satisfies the requirements of Definition 1.

Improper integral (9.2) can be interpreted geometrically in the following way. Let $f(x)$ be a continuous nonnegative function defined for $x \geqslant a$. Consider the domain Ω bounded below by

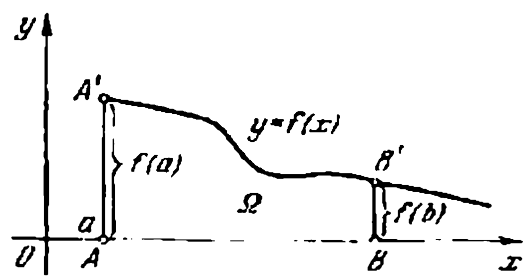


Fig. 9.1

the part $a \leqslant x < +\infty$ of the x-axis, above by the graph of the function and on the left by the line segment $x = a$, $0 \leqslant y \leqslant f(a)$. The definition of squarability and the notion of the area of a plane figure introduced in Sec. 3, § 1 of Chapter 1 are inapplicable to the domain Ω because it is unbounded. Let us take an arbitrary line segment $x = B > a$, $0 \leqslant y \leqslant f(B)$, which cuts off a squarable curvilinear trapezoid $ABB'A'$ (see Fig. 9.1) whose area is equal to the integral $\int\limits_a^B f(x)\,dx$. It is natural to extend the notion of squarability to the domain Ω if the area of the trapezoid $ABB'A'$ tends to a finite limit as $B \to +\infty$. In this case we say that Ω is *squarable* and call the limit of the area of the trapezoid $ABB'A'$ the *area of the domain* Ω. This area is expressed by improper integral (9.2).

The improper integral of the form

$$\int\limits_{-\infty}^{a} f(x)\,dx = \lim_{A\to-\infty}\int\limits_{A}^{a} f(x)\,dx \tag{9.3}$$

is defined by analogy with integral (9.2).

If both limits of integration are infinite we put, by definition,

$$\int\limits_{-\infty}^{+\infty} f(x)\,dx = \int\limits_{-\infty}^{a} f(x)\,dx + \int\limits_{a}^{+\infty} f(x)\,dx \tag{9.4}$$

where a is an arbitrary finite number, the integral $\int\limits_{-\infty}^{+\infty} f(x)\,dx$ *being regarded as* ***convergent*** *if and only if both integrals entering into the right-hand side of (9.4) converge.*

It can be easily shown that the integral $\int\limits_{-\infty}^{+\infty} f(x)\,dx$ is convergent or divergent irrespective of the particular choice of the point a, and if it is convergent its value does not depend on a.* Thus, an integral of the form $\int\limits_{-\infty}^{+\infty} f(x)\,dx$ is reduced to the integrals of the form $\int\limits_{a}^{+\infty} f(x)\,dx$ and $\int\limits_{-\infty}^{a} f(x)\,dx$. But an integral of

---

* The integral $\int\limits_{-\infty}^{+\infty} f(x)\,dx$ can be defined by the relation

$$\int\limits_{-\infty}^{+\infty} f(x)\,dx = \lim_{\substack{A\to-\infty\\ B\to+\infty}}\int\limits_{A}^{B} f(x)\,dx \tag{9.4'}$$

where A and B tend to their limits independently. In fact, by virtue of (9.2), (9.3) and (9.4) we have

$$\int\limits_{-\infty}^{+\infty} f(x)\,dx = \int\limits_{-\infty}^{a} f(x)\,dx + \int\limits_{a}^{+\infty} f(x)\,dx = \lim_{A\to-\infty}\int\limits_{A}^{a} f(x)\,dx +$$

$$+ \lim_{B\to+\infty}\int\limits_{a}^{B} f(x)\,dx = \lim_{\substack{A\to-\infty\\ B\to+\infty}}\int\limits_{A}^{B} f(x)\,dx$$

the last limit existing if and only if for $A\to-\infty$ and $B\to+\infty$ the corresponding limits $\lim\limits_{A\to-\infty}\int\limits_{A}^{a} f(x)\,dx$ and $\lim\limits_{B\to+\infty}\int\limits_{a}^{B} f(x)\,dx$ exist when A and B tend to their limits independently.

the type $\int\limits_{-\infty}^{a} f(x)\,dx$ is reduced to the corresponding integral of the form $\int\limits_{a}^{+\infty} f(x)\,dx$ by a simple change of variable if we substitute $-x$ for x, and therefore in what follows we shall limit ourselves to studying integrals of the form $\int\limits_{a}^{+\infty} f(x)\,dx$.

Let us consider some examples.

1. The integral $\int\limits_{0}^{+\infty} \sin x\,dx = \lim\limits_{B\to+\infty} \int\limits_{0}^{B} \sin x\,dx = \lim\limits_{B\to+\infty} (1-\cos B)$ is divergent since $\cos B$ has no limit and oscillates between -1 and $+1$ as $B\to+\infty$.

2. The integral $\int\limits_{-\infty}^{+\infty} \frac{dx}{1+x^2}$ converges because the finite limit

$$\lim_{\substack{B\to+\infty\\ A\to-\infty}} \int\limits_{A}^{B} \frac{dx}{1+x^2} = \lim_{\substack{B\to+\infty\\ A\to-\infty}} [\arctan B - \arctan A] = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) \quad \pi$$

exists.

3. The integral

$$\int\limits_{a}^{+\infty} \frac{C}{x^{\alpha}}\,dx \quad \text{where } C = \text{const} > 0 \text{ and } a > 0$$

is important for our further aims. It converges for $\alpha > 1$ and diverges for $\alpha \leqslant 1$. Indeed, we have

$$\int\limits_{a}^{B} \frac{C}{x^{\alpha}}\,dx = \begin{cases} C \ln \dfrac{B}{a} & \text{for } \alpha = 1 \\ C\,\dfrac{B^{1-\alpha} - a^{1-\alpha}}{1-\alpha} & \text{for } \alpha \neq 1 \end{cases}$$

and therefore

$$\lim_{B\to+\infty} \int\limits_{a}^{B} \frac{C}{x^{\alpha}}\,dx = \begin{cases} C\,\dfrac{a^{1-\alpha}}{\alpha-1} & \text{for } \alpha > 1 \\ +\infty & \text{for } \alpha \leqslant 1 \end{cases}$$

This integral is widely used when applying the comparison test to testing various improper integrals for convergence or divergence (see Theorem 9.3 in Sec. 4).

2. Reducing Improper Integral of the Form $\int\limits_{a}^{+\infty} f(x)\,dx$ to Nume-

rical Sequence and Numerical Series. Testing an improper integral $\int\limits_a^{+\infty} f(x)\,dx$ for convergence can be reduced to testing convergence of number sequences or number series.

According to Definition 1, the improper integral $\int\limits_a^{+\infty} f(x)\,dx$ is the limit of the function $F(B)=\int\limits_a^B f(x)\,dx$ for $B\to+\infty$. Applying to $F(B)$ the definition of the limit of a function in terms of sequences (e.g. see [8], Chapter 4, § 2) we arrive at the following test:

For the integral $\int\limits_a^{+\infty} f(x)\,dx$ to be convergent it is necessary and sufficient that for an arbitrary choice of a sequence of points

$$B_n > a,\quad n=1,2,\ldots;\quad B_n\to+\infty \quad \text{for } n\to\infty \tag{9.5}$$

the numerical sequence

$$\int\limits_a^{B_n} f(x)\,dx,\quad n=1,2,3,\ldots \tag{9.6}$$

converge to the same finite limit. In case the integral $\int\limits_a^{+\infty} f(x)\,dx$ is convergent the limit of sequence (9.6) is equal to the value of the integral.

Note that (9.6) is a sequence of partial sums of the series

$$\int\limits_a^{B_1} f(x)\,dx+\int\limits_{B_1}^{B_2} f(x)\,dx+\ldots+\int\limits_{B_{n-1}}^{B_n} f(x)\,dx+\ldots \tag{9.7}$$

and therefore the above test can be rephrased as follows:

For the integral $\int\limits_a^{+\infty} f(x)\,dx$ to be convergent it is necessary and sufficient that for any choice of sequence of points (9.5) *number series* (9.7) *converge and its sum be independent of the particular choice of the sequence. If the integral $\int\limits_a^{+\infty} f(x)\,dx$ converges the sum of series* (9.7) *is equal to the integral.*

Note 1. If $f(x)$ is a function with alternating sign on the interval $a\leqslant x<+\infty$, the convergence of series (9.7) for a specific choice of point sequence (9.5) does not imply, in the general case, the con-

vergence of the integral $\int\limits_a^{+\infty} f(x)\,dx$. For instance, the integral $\int\limits_0^{+\infty} \sin x\,dx$ is divergent (see Example 1 in Sec. 1) although the series $\sum\limits_{n=0}^{+\infty} \int\limits_{2\pi n}^{2\pi(n+1)} \sin x\,dx$ converges because all its terms are equal to zero.

If a function $f(x)$ retains its sign for all $x \geqslant a$, for instance, $f(x) \geqslant 0$ for all $x \geqslant a$, *then for the integral $\int\limits_a^{+\infty} f(x)\,dx$ to be convergent it is necessary and sufficient that series* (9.7) *converge for at least one choice of a monotone increasing sequence of type* (9.5).*

The necessity of this test follows from the above. Let us establish its sufficiency. Let $f(x) \geqslant 0$ for all $x \geqslant a$ and let series (9.7) converge for a monotone increasing sequence of type (9.5). The sequence (9.6) of partial sums of the series is monotone increasing (or nondecreasing) and tends to a finite limit J. We shall prove that for any other choice of a sequence

$$B'_m > a, \quad m = 1, 2, \ldots; \quad B'_m \to +\infty \quad \text{for } m \to +\infty \qquad (9.5')$$

the corresponding series

$$\int\limits_a^{B'_1} f(x)\,dx + \int\limits_{B'_1}^{B'_2} f(x)\,dx + \ldots + \int\limits_{B'_m}^{B'_{m+1}} f(x)\,dx + \ldots \qquad (9.7')$$

is converging and its sum is equal to J.

To prove the assertion we shall use partial sums of series (9.7) and (9.7′). Given $\varepsilon > 0$, there is B_{n_0} such that the inequality

$$J - \varepsilon < \int\limits_a^{B_{n_0}} f(x)\,dx < J$$

holds. Let us choose m_0 such that for all $m \geqslant m_0$ the inequality $B'_m \geqslant B_{n_0}$ is fulfilled. Furthermore, for any B'_m there exists $B_{n_m} > B'_m$, and therefore we have the inequality

$$J - \varepsilon < \int\limits_a^{B_{n_0}} f(x)\,dx \leqslant \int\limits_a^{B'_m} f(x)\,dx \leqslant \int\limits_a^{B_{n_m}} f(x)\,dx \leqslant J$$

* Compare with the well known Cauchy integral test for convergence of an infinite series (e.g. see [8], Chapter 13, § 2, Sec. 4).

for all $m \geqslant m_0$ since $f(x)$ is nonnegative. Consequently, $\lim\limits_{m\to+\infty} \int\limits_a^{B'_m} f(x)\,dx = J$, which is what we set out to prove.

Example. Take the function

$$f(x) = \begin{cases} 2^n & \text{for } n \leqslant x \leqslant n + \frac{1}{2^{2n}}, \quad n = 1, 2, \ldots \\ 0 & \text{for } n + \frac{1}{2^{2n}} < x < n+1, \; n = 1, 2, \ldots \end{cases}$$

Then the integral

$$\int\limits_1^{+\infty} f(x)\,dx = \sum_{n=1}^{+\infty} \int\limits_n^{n+1} f(x)\,dx = \sum_{n=1}^{+\infty} \frac{1}{2^n} = 1$$

is converging.

Note 2. The above example indicates that even if a function $f(x)$ is nonnegative the fact that the integral $\int\limits_a^{+\infty} f(x)\,dx$ is convergent does not imply that $f(x) \to 0$ for $x \to +\infty$.

3. Cauchy Criterion for Improper Integrals. The convergence of an improper integral

$$\int\limits_a^{+\infty} f(x)\,dx = \lim_{B\to+\infty} \int\limits_a^{B} f(x)\,dx \tag{9.2}$$

is equivalent to the existence of a finite limit of the function $F(B) = \int\limits_a^{B} f(x)\,dx$ for $B \to +\infty$. According to the Cauchy general criterion (e.g. see [8], Chapter 8, § 1, Sec. 2), $F(B)$ tends to a finite limit as $B \to +\infty$ if and only if for every $\varepsilon > 0$ there exists $B(\varepsilon)$ such that $|F(B'') - F(B')| < \varepsilon$ for all $B' > B(\varepsilon)$ and $B'' > B(\varepsilon)$. Taking the expression $\int\limits_a^{B} f(x)\,dx$ as the function $F(B)$ we obtain the following

Cauchy's Test (for Convergence of an Improper Integral of Type $\int\limits_a^{+\infty} f(x)\,dx$). *An integral $\int\limits_a^{+\infty} f(x)\,dx$ converges if and only if for every $\varepsilon > 0$ there exists $B(\varepsilon)$ such that the inequality*

$$\left| \int\limits_{B'}^{B''} f(x)\,dx \right| < \varepsilon \tag{9.8}$$

holds for all B', $B'' > B(\varepsilon)$, which is equivalent to the requirement that the integral

$$\int_{B'}^{B''} f(x)\,dx \tag{9.8'}$$

tends to zero for $B' \to +\infty$, $B'' \to +\infty$.

Cauchy test (9.8) may sometimes be directly applied to testing some integrals for convergence.

Example. Let us consider the integral $\int_0^{+\infty} \frac{\sin x}{x}\,dx$. Its integrand $f(x) = \frac{\sin x}{x}$ can be regarded as a continuous function if we put $f(x) = 1$ for $x = 0$. Integrating by parts we get

$$\int_{B'}^{B''} \frac{\sin x}{x}\,dx = \frac{\cos B'}{B'} - \frac{\cos B''}{B''} - \int_{B'}^{B''} \frac{\cos x}{x^2}\,dx$$

Therefore

$$\left| \int_{B'}^{B''} \frac{\sin x}{x}\,dx \right| \leqslant \frac{1}{B'} + \frac{1}{B''} + \left| \int_{B'}^{B''} \frac{|\cos x|}{x^2}\,dx \right| \leqslant$$

$$\leqslant \frac{1}{B'} + \frac{1}{B''} + \left| \int_{B'}^{B''} \frac{dx}{x^2} \right| \leqslant \frac{2}{B'} + \frac{2}{B''} \to 0 \text{ for } B',\ B'' \to +\infty$$

Hence, the integral $\int_0^{+\infty} \frac{\sin x}{x}\,dx$ is convergent.

But it should be noted that in many important applications the Cauchy test is more effectively used not for investigating concrete integrals but for establishing general tests providing some sufficient conditions for convergence which can be applied to practical problems. Before proceeding to study such tests we shall introduce the notion of an *absolutely convergent improper integral* which is analogous to the notion of an absolutely convergent numerical series.

4. Absolute Convergence. Tests for Absolute Convergence.

Definition 2. *Let a function $f(x)$ be integrable in the ordinary sense over every finite interval $a \leqslant x \leqslant B$, $a < B < +\infty$.* Improper*

---

* As is known, integrability of $f(x)$ in the ordinary sense over a finite interval implies integrability in the ordinary sense of $|f(x)|$ over the interval.

integral (9.2) *is said to be absolutely convergent if the integral*

$$\int_a^{+\infty} |f(x)|\,dx \tag{9.9}$$

converges.

Theorem 9.1. If integral (9.2) *converges absolutely it is convergent.*

Proof. Indeed, integral (9.10) being convergent, it follows that for every $\varepsilon > 0$ there is $B(\varepsilon)$ such that $\left|\int_{B'}^{B''} |f(x)|\,dx\right| < \varepsilon$ for all B', $B'' > B(\varepsilon)$. But we always have

$$\left|\int_{B'}^{B''} f(x)\,dx\right| \leqslant \left|\int_{B'}^{B''} |f(x)|\,dx\right| \tag{9.10}$$

which implies

$$\left|\int_{B'}^{B''} f(x)\,dx\right| \leqslant \left|\int_{B'}^{B''} |f(x)|\,dx\right| < \varepsilon \quad \text{for all } B',\ B'' > B(\varepsilon)$$

Hence the conditions of Cauchy's test are fulfilled for integral (9.2). Consequently, integral (9.2) is convergent and thus the theorem has been proved.

Note 1. The fact that integral (9.2) converges does not imply its absolute convergence. For example, the integral $\int_0^{+\infty} \frac{\sin x}{x}\,dx$ converges (see Sec. 3) while the integral $\int_0^{+\infty} \frac{|\sin x|}{x}\,dx$ is divergent. To prove the latter it is sufficient (see Sec. 2) to show that the number series $\sum\limits_{n=0}^{+\infty} \int_{\pi n}^{\pi(n+1)} \frac{|\sin x|}{x}\,dx$ is divergent which can easily be done by applying the comparison test for numerical series. In fact, we have

$$\int_{\pi n}^{\pi(n+1)} \frac{|\sin x|}{x}\,dx \geqslant \frac{1}{(n+1)\pi}\left|\int_{\pi n}^{\pi(n+1)} \sin x\,dx\right| = \frac{2}{(n+1)\pi} \quad \text{for } n \geqslant 1$$

and the series $\sum\limits_{n=1}^{+\infty} \frac{2}{\pi n} = \frac{2}{\pi} \sum\limits_{n=1}^{+\infty} \frac{1}{n}$ is divergent because it differs from the harmonic series only in the constant factor $\frac{2}{\pi}$.*

Note 2. Let $f(x)$ be a function defined for $a \leqslant x < +\infty$ and integrable over every finite interval $a \leqslant x \leqslant B$. Then, for any $a_1 > a$, we can assert that if the integral $\int\limits_{a_1}^{+\infty} f(x)\,dx$ is absolutely convergent the integral $\int\limits_{a}^{+\infty} f(x)\,dx$ is also absolutely convergent because for any of the integrals $\int\limits_{a}^{+\infty} |f(x)|\,dx$ and $\int\limits_{a_1}^{+\infty} |f(x)|\,dx$ to onvergent it is necessary and sufficient that

$$\int\limits_{B'}^{B''} |f(x)|\,dx \to 0 \quad \text{for } B',\ B'' \to +\infty$$

o test an improper integral for absolute convergence we usually apply comparison tests.

Theorem 9.2 (General Comparison Test). Suppose that the inequality

$$|f(x)| \leqslant g(x) \tag{9.11}$$

holds for all sufficiently large x. Then if the integral

$$\int\limits_{a}^{+\infty} g(x)\,dx \tag{9.12}$$

*is convergent the integral***

$$\int\limits_{a}^{+\infty} f(x)\,dx \tag{9.13}$$

is absolutely convergent.

---

* If an integral of the form $\int\limits_{a}^{+\infty} f(x)\,dx$ is convergent and the integral $\int\limits_{a}^{+\infty} |f(x)|\,dx$ is divergent the former is called conditionally convergent. Such integrals are treated in Sec. 5.

** In Theorems 9.2, 9.3, 9.3′, 9.3″ and 9.3‴ we suppose that the function $f(x)$ is integrable in the ordinary sense over every finite interval $a \leqslant x \leqslant B$, $a < B < +\infty$.

Proof. The conditions of Cauchy's test holding for convergent integral (9.1), inequality (9.11) implies that for any $\varepsilon > 0$ there exists $B(\varepsilon)$ such that

$$\left|\int_{B'}^{B''} |f(x)|\,dx\right| \leqslant \left|\int_{B'}^{B''} g(x)\,dx\right| < \varepsilon \quad \text{for all } B',\ B'' > B(\varepsilon)$$

i.e. the conditions of Cauchy's test also hold for the integral $\int\limits_a^{+\infty} |f(x)|\,dx$. Consequently, this integral is convergent whence it follows that integral (9.13) is absolutely convergent. The theorem has been proved.

In Example 3 of Sec. 1 we showed that

$$\int_a^{+\infty} \frac{C\,dx}{x^\alpha} = \lim_{B\to+\infty} \int_a^{B} \frac{C\,dx}{x^\alpha} = \begin{cases} C\dfrac{a^{1-\alpha}}{\alpha-1} & \text{for } \alpha > 1 \\ +\infty & \text{for } \alpha \leqslant 1 \end{cases}$$

for $a > 0$ and $C > 0$. Therefore, on the basis of the general comparison test we obtain

Theorem 9.3 **(*Special Comparison Test*)**. *Let an improper integral* $\int\limits_a^{+\infty} f(x)\,dx$ *be given.*

1. *If* $|f(x)| < \dfrac{C}{x^\alpha}$ *for all sufficiently large values of* x *where* $C \geqslant 0$ *and* $\alpha > 1$, *the integral converges absolutely.*

2. *If for all sufficiently large values of* x *the function* $f(x)$ *satisfies the inequality* $f(x) \geqslant \dfrac{C}{x^\alpha}$ *or* $f(x) \leqslant -\dfrac{C}{x^\alpha}$ *where* $C > 0$ *and* $\alpha \leqslant 1$, *the integral is divergent.*

Proof. (1) Putting $g(x) = \dfrac{C}{x^\alpha}$ in the general comparison test and taking into account that the integral $\int\limits_a^{+\infty} \dfrac{C\,dx}{x^\alpha} = \dfrac{Ca^{1-\alpha}}{\alpha-1}$ converges for $\alpha > 1$ and $a > 0$* we see that the integral $\int\limits_a^{+\infty} f(x)\,dx$ is absolutely convergent.

* We suppose that $a > 0$ because, if otherwise, a can be replaced by $a_1 > 0$, and the integrals $\int\limits_a^{+\infty} f(x)\,dx$ and $\int\limits_{a_1}^{+\infty} f(x)\,dx$ are simultaneously absolutely convergent or divergent.

(2) Let $f(x) \geqslant \frac{C}{x^\alpha}$ where $C>0$ and $\alpha \leqslant 1$ for all $x \geqslant a_1 > a$. Integrating between the limits a_1 and B we see that $\int\limits_{a_1}^{B} f(x)\,dx \geqslant$ $\geqslant C \int\limits_{a_1}^{B} \frac{dx}{x^\alpha} \to +\infty$ for $B \to +\infty$ since $\alpha \leqslant 1$. Hence, the integral $\int\limits_{a_1}^{+\infty} f(x)\,dx$ diverges, which implies that the integral $\int\limits_{a}^{+\infty} f(x)\,dx$ is also divergent.

If $f(x) \leqslant -\frac{C}{x^\alpha}$ for all $x \geqslant a_1 > a > 0$, $C>0$ and $\alpha \leqslant 1$ we can put $f^*(x) = -f(x)$ and then $f^*(x) \geqslant \frac{C}{x^\alpha}$ for all $x \geqslant a_1 > a > 0$, and consequently the integral $\int\limits_{a}^{+\infty} f^*(x)\,dx$ diverges. Therefore the integral

$$\int\limits_{a}^{+\infty} f(x)\,dx = \lim_{B\to+\infty} \int\limits_{a}^{B} f(x)\,dx = -\lim_{B\to+\infty} \int\limits_{a}^{B} f^*(x)\,dx$$

also diverges.

Note 1. Part 2 of Theorem 9.3 can be equivalently formulated as follows: *if a function $f(x)$ retains its sign and $|f(x)| \geqslant \frac{C}{x^\alpha}$ for all sufficiently large x $(x \geqslant a)$ where $C>0$ and $\alpha \leqslant 1$, the integral $\int\limits_{a}^{+\infty} f(x)\,dx$ diverges.*

Note 2. The fact that the condition $|f(x)| \geqslant \frac{C}{x^\alpha}$, $C>0$, $\alpha \leqslant 1$ holds for a function $f(x)$ for all sufficiently large $x \geqslant a$ may not imply that the integral $\int\limits_{a}^{+\infty} f(x)\,dx$ is divergent. This integral may turn out to be convergent if $f(x)$ is a function with alternating sign. For example, we can easily show that the integral $\int\limits_{1}^{+\infty} f(x)\,dx$ where the integrand is determined by the relation $f(x) =$ $= (-1)^{n+1} \frac{1}{x^\alpha}$, $n \leqslant x < n+1$, $n = 1, 2, 3, \ldots$, $0 < \alpha \leqslant 1$, conver-

ges although $|f(x)| = \frac{1}{x^\alpha}$ and the integral $\int_1^{+\infty} |f(x)|\,dx$ diverges for $0 < \alpha \leqslant 1$.

The test based on comparison with the function $\frac{C}{x^\alpha}$ can be rephrased as follows:

Theorem 9.3′ (Modified Special Comparison Test). Let the integrand in the integral $\int_a^{+\infty} f(x)\,dx$ *be representable, for all sufficiently large* x, *in the form* $f(x) = \frac{g(x)}{x^\alpha}$. *Then* (1) *if* $g(x)$ *is bounded in its modulus and* $\alpha > 1$, *the integral is absolutely convergent,* (2) *if* $g(x)$ *retains its sign,* $|g(x)| \geqslant \text{const} > 0$ *and* $\alpha \leqslant 1$ *the integral diverges.*

The following form of the comparison test based on comparison with the function $\frac{C}{x^\alpha}$ sometimes proves to be convenient:

Theorem 9.3″ (the Limiting Form of Special Comparison Test). Let the limit $\lim\limits_{x\to+\infty} |f(x)|\,x^\alpha = C$ *exist. Then* (1) *if* $0 \leqslant C < +\infty$ *and* $\alpha > 1$ *the integral* $\int_a^{+\infty} f(x)\,dx$ *is absolutely convergent, and* (2) *if* $0 < C \leqslant +\infty$, $\alpha \leqslant 1$ *and* $f(x)$ *retains its sign for all sufficiently large* x, *the integral is divergent.*

Proof. (1) If $0 \leqslant C < +\infty$, we have, for all sufficiently large x, the inequality

$$|f(x)|\,x^\alpha \leqslant 2C, \text{ i.e. } |f(x)| \leqslant \frac{2C}{x^\alpha}, \text{ for } C > 0$$

and the inequality

$$|f(x)|\,x^\alpha \leqslant 1, \text{ i.e. } |f(x)| \leqslant \frac{1}{x^\alpha}, \quad \text{for } C = 0$$

Therefore, by Theorem 9.3. the integral $\int_a^{+\infty} f(x)\,dx$ converges absolutely.

(2) If $0 < C \leqslant +\infty$ and $\alpha \leqslant 1$, we have

$$|f(x)|\,x^\alpha > \frac{C}{2}, \text{ i.e. } |f(x)| > \frac{C}{2x^\alpha}, \text{ for } C < +\infty$$

and

$$|f(x)|\,x^{\alpha} > 1, \quad \text{i.e.} \quad |f(x)| > \frac{1}{x^{\alpha}}, \quad \text{for } C = +\infty$$

for all sufficiently large x, and hence, on the basis of Note 1 after Theorem 9.3′, we conclude that the integral $\int\limits_a^{+\infty} f(x)\,dx$ diverges. The theorem has been proved.

Note 3. Theorem 9.3″ (the modified comparison test in the limit form) embraces a narrower class of functions than Theorem 9.3′ (the modified comparison test) because Theorem 9.3″, unlike Theorem 9.3′ is only applicable if a finite or infinite limit of $|f(x)|\,x^{\alpha}$ exists for $x \to +\infty$.

Theorem 9.3″ implies

Theorem 9.3‴ (Special Comparison Test in Terms of Orders of Infinitesimals). *Let $|f(x)|$ be an infinitesimal of the order of $\frac{1}{x^{\alpha}}$ for $x \to +\infty$. Then (1) if $\alpha > 1$ the integral $\int\limits_a^{+\infty} f(x)\,dx$ converges absolutely and (2) if $\alpha \leqslant 1$ and $f(x)$ retains its sign for all sufficiently large x, the integral $\int\limits_a^{+\infty} f(x)\,dx$ diverges.*

We remind the reader that $f(x)$ is said to be an *infinitesimal of the order of* $\frac{1}{x^{\alpha}}$ $(\alpha > 0)$ for $x \to +\infty$ if

$$\lim_{x \to +\infty} \frac{|f(x)|}{\left(\frac{1}{x^{\alpha}}\right)} = \lim_{x \to +\infty} |f(x)|\,x^{\alpha} = C \quad \text{where } 0 < C < +\infty$$

Note 4. It is obvious that Theorem 9.3‴ is applicable to a still narrower class of functions than Theorem 9.3″ since its conditions include the existence of a finite limit of $|f(x)|\,x^{\alpha}$, for $x \to +\infty$, different from zero and infinity.

Examples

1. By Theorem 9.3, the **Euler-Poisson* integral** $\int\limits_0^{+\infty} e^{-x^2}\,dx$ is convergent since the exponential function e^{-x^2} decreases faster than any negative power of x as $x \to +\infty$ and, consequently, we have

$$e^{-x^2} < \frac{C}{x^2}$$

* Poisson, Siméon Denis (1781-1840), a French mathematician

for all sufficiently large x where $C = \text{const} > 0$ (here we have put $\alpha = 2$ but any other number exceeding 1 can also be taken as α). Theorem 9.3′ also indicates that this integral is convergent because we have

$$\lim_{x\to+\infty} x^2e^{-x^2} = 0 \quad (\alpha = 2)$$

2. The integral $\int_1^{+\infty} e^{-x}x^{p-1}\,dx$ converges for all real values of p. Indeed, to prove this we can apply, for instance, Theorem 9.3′ since the relation

$$\lim_{x\to+\infty} x^2e^{-x}x^{p-1} = 0$$

is fulfilled for all such p.

3. Let us consider the integral $\int_1^{+\infty} \frac{x^m}{1+x^n}\,dx$ for $n \geqslant 0$. We have

$$\frac{x^m}{1+x^n} = \frac{x^m}{x^n}\,\frac{1}{1+x^{-n}} = \frac{g(x)}{x^{n-m}} \qquad g(x) = \frac{1}{1+x^{-n}}$$

where $\frac{1}{2} \leqslant g(x) \leqslant 1$ for $x \geqslant 1$. Therefore, by Theorem 9.3′, the integral is convergent for $n - m > 1$ and divergent for $n - m \leqslant 1$.

5. **Conditional Convergence.**

Definition 3. An integral

$$\int_a^{+\infty} f(x)\,dx \tag{9.14}$$

is said to be **conditionally convergent** *if it converges while the integral*

$$\int_a^{+\infty} |f(x)|\,dx \tag{9.15}$$

diverges.

The Abel theorem given below makes it possible to test some conditionally convergent integrals for convergence.

Theorem 9.4 (Abel's Test). Let $\varphi(x)$ be continuous and $g(x)$ be continuously differentiable on the infinite interval $a \leqslant x < +\infty$. If the antiderivative $\Phi(B) = \int_a^B \varphi(x)\,dx$ is bounded on the interval $a \leqslant B < +\infty$ and $g(x)$ is a monotone decreasing function which

tends to zero for $x \to +\infty$, the integral

$$\int_{a}^{+\infty} \varphi(x)\, g(x)\, dx \tag{9.16}$$

is convergent.

Proof. We shall show that, under the conditions of the theorem, the requirements of Cauchy's test are fulfilled for integral (9.16). Integrating by parts we obtain

$$\int_{B'}^{B''} \varphi(x)\, g(x)\, dx = \Phi(B'')\, g(B'') - \Phi(B')\, g(B') - \int_{B'}^{B''} \Phi(x)\, g'(x)\, dx$$

The function $g(x)$ being monotone and decreasing as $x \to +\infty$, we have $g'(x) \leqslant 0$, and it is therefore allowable to apply the generalized mean value theorem to the last integral. This results in

$$\int_{B'}^{B''} \Phi(x)\, g'(x)\, dx = \Phi(\xi) \int_{B'}^{B''} g'(x)\, dx = \Phi(\xi)\, [g(B'') - g(B')]$$

where ξ lies between B' and B''. Consequently,

$$\int_{B'}^{B''} \varphi(x)\, g(x)\, dx = \Phi(B'')\, g(B'') - \Phi(B')\, g(B') - \Phi(\xi)\, [g(B'') - g(B')]$$

The antiderivative $\Phi(B)$ being bounded, it follows that

$$\int_{B'}^{B''} \varphi(x)\, g(x)\, dx \to 0 \quad \text{for } B',\ B'' \to +\infty$$

since $g(B)$ tends to zero as $B \to +\infty$. The theorem has thus been proved.

Examples

1. The integral $\int_{\pi}^{+\infty} \frac{\sin x}{x^{\alpha}}\, dx$ is convergent for $\alpha > 0$ because if we put $\varphi(x) = \sin x$ and $g(x) = \frac{1}{x^{\alpha}}$, we have

$$|\Phi(x)| = \left| \int_{\pi}^{x} \varphi(x)\, dx \right| = \left| \int_{\pi}^{x} \sin x\, dx \right| = |\cos \pi - \cos x| \leqslant 2$$

for $\pi \leqslant x \leqslant +\infty$, and $g(x) = \frac{1}{x^{\alpha}}$ is a monotone decreasing function tending to zero for $x \to +\infty$ and $\alpha > 0$.

Note. We can prove that the above integral is convergent without resorting to Abel's test if we apply Cauchy's test and integration by parts as it was done at the end of Sec. 4.

2. Taking the integral

$$\int_{e}^{+\infty} \frac{(\ln x)\sin x}{x}\,dx$$

and putting $\varphi(x) = \sin x$ and $g(x) = \frac{\ln x}{x}$ we see that, by Abel's test, it is convergent.

3. Let us consider the Fresnel* integral $\int_{0}^{+\infty} \sin(x^2)\,dx$ which is used in optics. Putting $x^2 = t$ we obtain

$$\int_{0}^{+\infty} \sin(x^2)\,dx = \int_{0}^{+\infty} \frac{\sin t}{2\sqrt{t}}\,dt$$

The Abel test indicates that the latter integral converges.

6. Extending Methods of Evaluating Integrals to the Case of Improper Integrals. When evaluating improper integrals we can change variables, integrate by parts and represent the integral of a sum of functions as the sum of the integrals of these functions, that is apply all the methods of evaluating proper integrals. The corresponding formulas are valid provided all the integrals entering into them are convergent.

Let us illustrate the significance of the latter requirement by taking a concrete example. The integral $\int_{3}^{+\infty} \frac{dx}{x^2+x-2}$ is convergent (which, for instance, can be established on the basis of the special comparison test). Let us write down the decomposition of the integrand into partial fractions (e.g. see [8], Chapter 7, § 7):

$$\frac{1}{x^2+x-2} = -\frac{1}{3(x+2)} + \frac{1}{3(x-1)} \qquad (*)$$

It is obvious that the integrals $\int_{3}^{+\infty} \frac{dx}{x+2}$ and $\int_{3}^{+\infty} \frac{dx}{x-1}$ are divergent. Therefore the equality

$$\int_{3}^{+\infty} \frac{dx}{x^2+x-2} = -\frac{1}{3}\int_{3}^{+\infty} \frac{dx}{x+2} + \frac{1}{3}\int_{3}^{+\infty} \frac{dx}{x-1}$$

* Fresnel, Augustin Jean (1788-1827), a French optician and geometer.

is incorrect. To use decomposition (*) for evaluating the integral we can integrate (*) from 0 to A and then transform the right-hand side of the resulting equality to the form

$$\int_3^A \frac{dx}{x^2+x-2} = -\frac{1}{3}\int_3^A \frac{dx}{x+2} + \frac{1}{3}\int_3^A \frac{dx}{x-1} = \frac{1}{3}\ln\left[\frac{5}{2}\,\frac{A-1}{A+2}\right]$$

Now passing to the limit in the last equality as $A \to +\infty$ we obtain

$$\int_3^{+\infty} \frac{dx}{x^2+x-2} = \frac{1}{3}\ln\frac{5}{2}$$

§ 2. INTEGRALS OF UNBOUNDED FUNCTIONS WITH FINITE AND INFINITE LIMITS OF INTEGRATION

We shall first consider the integrals with finite limits of integration. Let a function $f(x)$ be defined everywhere on an interval $[a, b]$

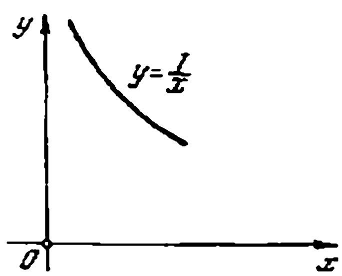


Fig. 9.2

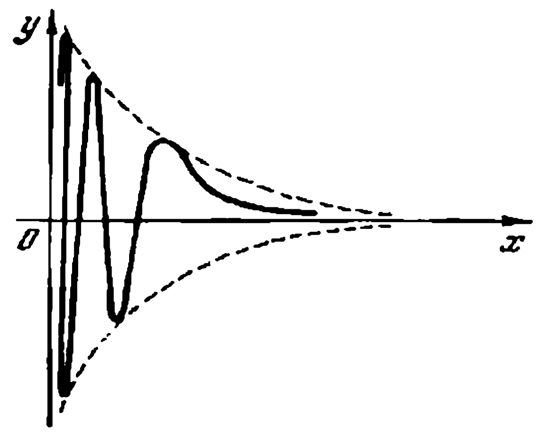


Fig. 9.3

except possibly at a finite number of points. A point $x_0 \in [a, b]$ will be referred to as a **singular point** of $f(x)$ if $f(x)$ is unbounded in every neighbourhood of the point x_0. For example, the function

$$f(x) = \begin{cases} \dfrac{1}{x} & \text{for } 0 < x \leqslant 1 \\ 0 & \text{for } x = 0 \end{cases}$$

has the singular point $x = 0$ (see Fig. 9.2).

The point $x = 0$ is also a singular point for the function

$$f(x) = \begin{cases} \dfrac{1}{x}\sin\dfrac{1}{x} & \text{for } 0 < x \leqslant 1 \\ 0 & \text{for } x = 0 \end{cases}$$

(Fig. 9.3). It should be noted that in the last example the function $f(x)$ does not have an infinite limit when $x \to 0$ because it turns

into zero infinitely many times in every arbitrarily small neighbourhood of the point $x = 0$.

Definition. Let a function $f(x)$ be defined on an interval $[a, b]$ everywhere except possibly at a finite number of points. If $x = b$ is a singular point of the function $f(x)$ and the integral $\int\limits_a^{b-\mu} f(x)\,dx$ exists for every μ, $0 < \mu < b - a$, the improper integral $\int\limits_a^b f(x)\,dx$ is understood as the limit

$$\int\limits_a^b f(x)\,dx = \lim_{\mu\to+0} \int\limits_a^{b-\mu} f(x)\,dx \tag{9.17}$$

If this finite limit exists integral (9.17) is said to be convergent, if otherwise it is called divergent.

We similarly define the improper integral in the case when the singular point of $f(x)$ coincides with the left end point $x = a$ of the interval of integration $[a, b]$:

$$\int\limits_a^b f(x)\,dx = \lim_{\lambda\to+0} \int\limits_{a+\lambda}^{b} f(x)\,dx \tag{9.18}$$

If both end points $x=a$ and $x=b$ are singular points of $f(x)$ the integral is defined as

$$\int\limits_a^b f(x)\,dx = \lim_{\substack{\lambda\to+0\\ \mu\to+0}} \int\limits_{a+\lambda}^{b-\mu} f(x)\,dx \tag{9.19}$$

If an interior point $x=c$, $a<c<b$, of the interval $[a, b]$ is a singular point of $f(x)$ we put

$$\int\limits_a^b f(x)\,dx = \lim_{\substack{\lambda\to+0\\ \mu\to+0}} \left[\int\limits_a^{c-\mu} f(x)\,dx + \int\limits_{c+\lambda}^{b} f(x)\,dx\right] \tag{9.20}$$

Let us now discuss the conditions guaranteeing the convergence of an improper integral of an unbounded function. Applying Cauchy's criterion to the function

$$F(\mu) = \int\limits_a^{b-\mu} f(x)\,dx \tag{9.21}$$

for $\mu \to +0$ we obtain

Cauchy's Test (for Improper Integral (9.17)*). For integral* (9.17) *to be convergent it is necessary and sufficient that for every* $\varepsilon > 0$ *there exist* $\delta = \delta(\varepsilon) > 0$ *such that*

$$\left| \int_{b-\mu'}^{b-\mu''} f(x)\,dx \right| < \varepsilon \quad \text{for all } 0 < \mu',\ \mu'' < \delta(\varepsilon)$$

Cauchy's test for integrals (9.18)-(9.20) can be formulated in a similar manner. It can easily be shown that integrals (9.19) and (9.20) converge if and only if both integrals $\int_a^c f(x)\,dx$ and $\int_c^b f(x)\,dx$ are convergent, and that in the case of integral (9.19) the point c can be chosen arbitrarily. If integrals (9.19) and (9.20) are convergent we have the equality

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx \tag{9.22}$$

for both integrals.

The same idea can be used for defining an improper integral with an integrand having a finite number of singular points on its interval of integration $[a, b]$: the interval is broken up into a finite number of subintervals in such a way that the function $f(x)$ has a singular point at only one of the end points of each subinterval.

Thus, the general case is reduced to integrals (9.17) and (9.18). But integral (9.18) reduces to the corresponding integral (9.17) if we substitute $-x$ for x. Therefore we shall limit ourselves to studying integrals of form (9.17).

Absolute convergence and conditional convergence are defined for integrals of unbounded functions like in the case of integrals with infinite limits of integrations. Cauchy's test makes it possible to prove that an absolutely convergent integral is convergent and also to establish the following test.

General Comparison Test. If b is the only singular point of $f(x)$ on $[a, b]$ and $|f(x)| \leqslant g(x)$ for all $x \in [a, b)$ lying sufficiently close to b and if the integral $\int_a^b g(x)\,dx$ converges, the integral $\int_a^b f(x)\,dx$ is absolutely convergent.

Let us now formulate a test based on comparison with the function $\frac{C}{(b-x)^{\alpha}}$ which is analogous to Theorem 9.3.

Special Comparison Test. Let a function $f(x)$ defined on $[a, b]$ have a singular point at the end point $x=b$ and let the integral $\int_a^{b-\mu} f(x)\,dx$ exist for every μ, $0<\mu<b-a$. Then

(1) *if we have*

$$|f(x)|\leqslant\frac{C}{(b-x)^{\alpha}}\quad where\ 0\leqslant C<+\infty,\ \alpha<1 \tag{9.23}$$

for all $x\in[a, b)$ lying sufficiently close to b, the integral $\int_a^b f(x)\,dx$ converges absolutely;

(2) *if we have*

$$f(x)>\frac{C}{(b-x)^{\alpha}}\quad where\ C>0,\quad \alpha\geqslant 1 \tag{9.24}$$

for all $x\in[a, b)$ which are sufficiently close to b or

$$f(x)<-\frac{C}{(b-x)^{\alpha}}\quad where\ C>0,\quad \alpha\geqslant 1 \tag{9.24'}$$

for all $x\in[a, b)$ lying sufficiently close to b, this integral diverges.

Proof. (1) In this case we have

$$\left|\int_{b-\mu'}^{b-\mu''} f(x)\,dx\right|\leqslant\left|\int_{b-\mu'}^{b-\mu''} |f(x)|\,dx\right|\leqslant C\left|\int_{b-\mu'}^{b-\mu''}\frac{dx}{(b-x)^{\alpha}}\right|=$$

$$=C\left|\frac{(\mu')^{1-\alpha}-(\mu'')^{1-\alpha}}{1-\alpha}\right|\to 0$$

for $\alpha<1$ and μ', $\mu''\to 0$. Consequently, the integrals $\int_a^b f(x)\,dx$ and $\int_a^b |f(x)|\,dx$ are convergent.

(2) Let us suppose that $f(x)$ is nonnegative.* Then we have

$$f(x)>\frac{C}{(b-x)^{\alpha}}\ \text{for}\ a<a_1\leqslant x<b,\ \alpha\geqslant 1$$

* If $f(x)$ is nonpositive we introduce the function $f^*(x)=-f(x)$ which is nonnegative, and if the integral $\int_a^b f^*(x)\,dx$ diverges the integral $\int_a^b f(x)\,dx=-\int_a^b f^*(x)\,dx$ also diverges.

and

$$\int_{a_1}^{b-\mu} f(x)\,dx > \int_{a_1}^{b-\mu} \frac{C}{(b-x)^{\alpha}}\,dx \to +\infty \text{ for } \mu \to 0+0 \text{ and } \alpha \geqslant 1$$

because

$$\int_{a_1}^{b-\mu} \frac{C}{(b-x)^{\alpha}}\,dx = \begin{cases} C\left[\dfrac{\mu^{1-\alpha}}{1-\alpha} - \dfrac{(b-a_1)^{1-\alpha}}{1-\alpha}\right] & \text{for } \alpha > 1 \\ C \ln \dfrac{b-a_1}{\mu} & \text{for } \alpha = 1 \end{cases}$$

Hence, the integral $\int_{a_1}^{b} f(x)\,dx$ diverges and thus the integral $\int_{a} f(x)\,dx$ also diverges

Note. Part (2) of the above theorem can be formulated equivalently in the following form: *if* $|f(x)| \geqslant \frac{C}{(x-b)^{\alpha}}$ *for all* x *lying sufficiently close to* b *where* $C > 0$, $\alpha \geqslant 1$ *and* $f(x)$ *retains its sign for these values of* x, *the integral* $\int_a^b f(x)\,dx$ *is divergent.*

The above special comparison test can be rephrased as follows:

Modified Special Comparison Test. *Let a function* $f(x)$ *be integrable in the ordinary sense on every finite interval* $a \leqslant x \leqslant b - \lambda$ *where* $0 < \lambda < b - a$. *Suppose that this function can be represented in the form* $f(x) = \frac{g(x)}{(b-x)^{\alpha}}$ *in a neighbourhood of* b *(i.e. for* $b - \delta < x < b$, $\delta > 0$, $\delta < b - a$). *Then*

(1) *if* $g(x)$ *is bounded in its modulus and* $\alpha < 1$, *the integral* $\int_a^b f(x)\,dx$ *converges absolutely*;

(2) *if* $g(x)$ *retains its sign in a neighbourhood of* b, $|g(x)| \geqslant \text{const} > 0$ *and* $\alpha \geqslant 1$, *the integral* $\int_a^b f(x)\,dx$ *diverges.*

The modified special comparison test can be formulated in like manner in the case when the only singular point of $f(x)$ on the interval $[a, b]$ is located at the end point $x = a$.

It is also possible to formulate and prove the special comparison test in the limiting form, which we leave to the reader.

We remind the reader that $|f(x)|$ is said to be *of the order of* $\frac{1}{(b-x)^{\alpha}}$ *for* $x \to b-0$ if

$$\lim_{x\to b-0} \frac{|f(x)|}{\left(\frac{1}{(b-x)^{\alpha}}\right)} = \lim_{x\to b-0} (b-x)^{\alpha} \ |f(x)| = C \quad \text{where} \quad 0 < C < +\infty$$

Let us now formulate

Special Comparison Test in Terms of Orders of Infinities. *Let* $|f(x)|$ *be an infinitely large quantity of the order of* $\frac{1}{(b-x)^{\alpha}}$ $(\alpha > 0)$ *for* $x \to b-0$.* *Then*

(1) *if* $\alpha < 1$ *the integral* $\int\limits_a^b f(x)\,dx$ *is absolutely convergent*;

(2) *if* $\alpha \geqslant 1$ *and* $f(x)$ *retains its sign in a neighbourhood of* $x = b$ (*i.e. for* $b - \lambda < x < b$, $0 < \lambda < b - a$) *this integral is divergent.*

This test is formulated similarly when $f(x)$ has a singular point not at the end point $x = b$ but at the end point $x = a$ of the interval $[a, b]$.

Examples

1. The integral $\int\limits_0^1 \frac{dx}{\sqrt{1-x^3}}$ converges since we have

$$f(x) = \frac{1}{\sqrt{1-x^3}} = \frac{1}{(1-x)^{1/2}} \frac{1}{(1+x+x^2)^{1/2}} = \frac{1}{(1-x)^{1/2}} g(x)$$

where $g(x) = \frac{1}{(1+x+x^2)^{1/2}}$ is a bounded function. Here we have $a = 0$, $b = 1$ and $\alpha = \frac{1}{2}$.

2. Consider the integral $\int\limits_0^1 \frac{\sin x}{x^p}\,dx$. We have

$$f(x) = \frac{\sin x}{x^p} = \frac{1}{x^{p-1}} \frac{\sin x}{x} = \frac{1}{x^{p-1}} g(x)$$

where $g(x) = \frac{\sin x}{x}$ is a function bounded in its modulus, and $\sin x \leqslant g(x) \leqslant 1$. Here we have $a = 0$, $b = 1$ and $\alpha = p - 1$. Therefore the integral converges for $\alpha = p - 1 < 1$ and diverges for $\alpha = p - 1 \geqslant 1$. Thus, the integral $\int\limits_0^1 \frac{\sin x}{x^p}\,dx$ is convergent for $p < 2$ and divergent for $p \geqslant 2$.

---

* The function $f(x)$ is supposed to be integrable in the ordinary sense over every interval $a \leqslant x \leqslant b - \lambda$, $0 < \lambda < b - a$.

Abel's test for improper integrals with finite limits of integration can be formulated and proved by analogy with Abel's test for improper integrals with infinite limits of integration (see § 1, Sec. 5) and we leave this to the reader.

Finally, we can change variables, integrate by parts and break up integrals into sums of integrals in the case of improper integrals with finite limits of integration under the same conditions as in the case of improper integrals with infinite limits of integration (see § 1, Sec. 6).

Now we shall briefly discuss the integrals with infinite limits of integration of unbounded functions having a finite number of singular points. If an improper integral is taken over an interval $a \leqslant x < +\infty$ or $-\infty < x \leqslant a$ or over the whole x-axis $-\infty < < x < +\infty$ we break up the interval of integration by means of one or two points of division into one finite interval containing all the singular points of the integrand $f(x)$ and one or two semi-infinite intervals without singular points of $f(x)$. This reduces the improper integral of the general type to the above special cases. The original integral is, by definition, understood as being equal to the sum of the integrals taken over the subintervals the original interval of integration is broken into.

The original integral is said to be convergent if and only if all the integrals over the subintervals are convergent. If at least one of these integrals is divergent, the original integral is regarded as divergent.

It can be shown that the definition of convergence of the original integral and its numerical value (provided this integral is convergent) are independent of the choice of the points of division.

Examples

3. Take the integral $\int\limits_0^{+\infty} e^{-x}x^{p-1}\,dx$. If $p-1<0$, the integrand has a singular point at $x=0$. Therefore let us divide the interval of integration into two intervals, e.g. $[0, 1]$ and $(1, +\infty)$, by means of the point $x=1$. This results in

$$\int\limits_0^{+\infty} e^{-x}x^{p-1}\,dx = \int\limits_0^{1} e^{-x}x^{p-1}\,dx + \int\limits_1^{+\infty} e^{-x}x^{p-1}\,dx$$

The integral $\int\limits_0^{1} e^{-x}x^{p-1}\,dx = \int\limits_0^{1} \frac{e^{-x}}{x^{1-p}}\,dx$ converges for $1-p<1$, that is for $p>0$, and diverges for $p \leqslant 0$. As has been shown (see § 1, Sec. 4), the integral $\int\limits_1^{+\infty} e^{-x}x^{p-1}\,dx$ converges for all values of p,

$-\infty < p < +\infty$, and, consequently, the integral $\int\limits_0^{+\infty} e^{-x}x^{p-1}\,dx$*
converges for all $p > 0$ and diverges for all $p \leqslant 0$.

4. Consider the integral $\int\limits_0^1 x^p \ln^q \frac{1}{x}\,dx$. Performing the substitution $\ln\frac{1}{x} = t$ $\left(\frac{1}{x} = e^t,\ x = e^{-t}\right.$ and $\left.dx = -e^{-t}\,dt\right)$ we obtain

$$\int\limits_0^1 x^p \ln^q \frac{1}{x}\,dx = -\int\limits_{+\infty}^0 e^{-pt}t^q e^{-t}\,dt = \int\limits_0^{+\infty} t^q e^{-(p+1)t}\,dt$$

The integrand in the last integral has a singular point $t = 0$ and therefore we break up the integral into two integrals:

$$\int\limits_0^{+\infty} t^q e^{-(p+1)t}\,dt = \int\limits_0^1 e^{-(p+1)t}t^q\,dt + \int\limits_1^{+\infty} e^{-(p+1)t}t^q\,dt$$

The integral $\int\limits_0^1 e^{-(p+1)t}t^q\,dt = \int\limits_0^1 \frac{e^{-(p+1)t}}{t^{-q}}\,dt$ is convergent only for $-q < 1$, i.e. for $q > -1$, irrespective of the values of p. The integral $\int\limits_1^{+\infty} e^{-(p+1)t}t^q\,dt$ (with $q > -1$) converges only if $p+1 > 0$, that is only if $p > -1$. Consequently, the integral $\int\limits_0^1 x^p \ln^q \frac{1}{x}\,dx$ converges for $p > -1$ and $q > -1$ and diverges for all the other values of p and q.

5. Taking the integral $\int\limits_e^{+\infty} \frac{dx}{x^p(\ln x)^q(\ln\ln x)^r}$ and performing the change of variable $\ln\ln x = t$ we find that the integral converges for $p > 1$ (and arbitrary q) only if $r < 1$. If $p = 1$ it converges only if $r < 1$ and $q > 0$. Finally, if $p < 1$ the integral diverges for any r and q.

---

* This integral is known as Euler's integral of the second kind (the gamma function) and is designated by $\Gamma(p)$, i.e. $\Gamma(p) = \int\limits_0^{+\infty} e^{-x}x^{p-1}\,dx$ (see § 3 of Chapter 10).

§ 3. CAUCHY'S PRINCIPAL VALUE OF A DIVERGENT IMPROPER INTEGRAL

Let a function $f(x)$ be integrable in the ordinary sense on every finite interval of the x-axis. If the limit $\lim\limits_{\substack{A\to-\infty\\B\to+\infty}}\int\limits_A^B f(x)\,dx$ does not exist when A and B tend independently to their limits, i.e. the integral $\int\limits_{-\infty}^{+\infty} f(x)\,dx$ diverges, but the limit $\lim\limits_{A\to+\infty}\int\limits_{-A}^{A} f(x)\,dx$ exists, the latter is called the **(Cauchy) principal value** of the divergent integral $\int\limits_{-\infty}^{+\infty} f(x)\,dx$. In this case we write

$$\text{v.p.}\int\limits_{-\infty}^{+\infty} f(x)\,dx=\lim_{A\to+\infty}\int\limits_{-A}^{A} f(x)\,dx \tag{9.25}$$

(the notation v.p. originates from French *valeur principale* principal value).

Now let $f(x)$ be a function defined on an interval $[a, b]$ and having only one singular point c, $a<c<b$. Suppose that the integral

$$\int\limits_a^b f(x)\,dx=\lim_{\substack{\lambda\to0+0\\ \mu\to0+0}}\left\{\int\limits_a^{c-\lambda} f(x)\,dx+\int\limits_{c+\mu}^{b} f(x)\,dx\right\}$$

is divergent, that is the limit of the expression in curly brackets does not exist when $\lambda>0$ and $\mu>0$ independently tend to zero. Then if the limit of this expression exists for $\lambda=\mu\to0+0$ it is called the **principal value** of the divergent integral $\int\limits_a^b f(x)\,dx$ and denoted as

$$\text{v.p.}\int\limits_a^b f(x)\,dx=\lim_{\lambda\to0+0}\left\{\int\limits_a^{c-\lambda} f(x)\,dx+\int\limits_{c+\lambda}^{b} f(x)\,dx\right\} \tag{9.26}$$

Examples

1. If $f(x)$ is an *odd function* (i.e. $f(-x)\equiv-f(x)$; see Sec. 6 in § 1 of Chapter 11) defined on $(-\infty, +\infty)$ the principal value

$$\text{v.p.}\int\limits_{-\infty}^{+\infty} f(x)\,dx=\lim_{A\to+\infty}\int\limits_{-A}^{A} f(x)\,dx=0$$

always exists.

2. If $f(x)$ is an *even function* on $(-\infty, +\infty)$ (i.e. $f(-x) \equiv f(x)$; see Sec. 6 in § 1 of Chapter 11) we have

$$\int_{-A}^{A} f(x)\,dx = 2\int_{0}^{A} f(x)\,dx = 2\int_{-A}^{0} f(x)\,dx$$

for any A. Therefore, if the integral $\int_{-\infty}^{+\infty} f(x)\,dx$ of an even function diverges, i.e. if at least one of the integrals $\int_{-\infty}^{0} f(x)\,dx$ and $\int_{0}^{+\infty} f(x)\,dx$ does not exist, the principal value v.p. $\int_{-\infty}^{+\infty} f(x)\,dx$ does not exist either.

3. Let us apply the notion of the principal value of a divergent improper integral to evaluating the integral

$$2J = \int_{-\infty}^{+\infty} \frac{x^{2m}}{1+x^{2n}}\,dx, \text{ where } n \text{ and } m \text{ are integers and } 0 < m < n \quad (9.27)$$

This integral plays an important role in the theory of *Euler's integrals* (see § 3 of Chapter 10). We have

$$\left|\frac{x^{2m}}{1+x^{2n}}\right| < \frac{C}{x^2} \text{ for } x \to \pm\infty, \text{ where } C = \text{const} > 0$$

Besides, all the roots $x_k = e^{i\frac{(2k+1)\pi}{2n}} = a_k + ib_k$, $k = 0, 1, \ldots, 2n-1$, of the equation $1 + x^{2n} = 0$ are not real. Therefore the integrand has no singular points on the x-axis and the integral thus converges. Taking the decomposition of the rational fraction $\frac{x^{2m}}{1+x^{2n}}$ into partial fractions and integrating from $-l$ to l $(l > 0)$ we find*

$$\int_{-l}^{l} \frac{x^{2m}}{1+x^{2n}}\,dx = \sum_{k=0}^{2n-1} A_k \int_{-l}^{l} \frac{dx}{x-x_k} = \sum_{k=0}^{2n-1} A_k \int_{-l}^{l} \frac{dx}{(x-a_k)-ib_k} =$$

$$= \sum_{k=0}^{2n-1} A_k \left\{ \int_{-l}^{l} \frac{x-a_k}{(x-a_k)^2+b_k^2}\,dx + i\int_{-l}^{l} \frac{b_k}{(x-a_k)^2+b_k^2}\,dx \right\} =$$

---

* An integral of a complex function $u(x)+iv(x)$ of a real variable x where $u(x)$ and $v(x)$ are real functions is defined as $\int [u(x)+iv(x)]\,dx = \int u(x)\,dx + i\int v(x)\,dx$.

$$= \sum_{k=0}^{2n-1} A_k \left\{ \ln \frac{(l-a_k)^2+b_k^2}{(l+a_k)^2+b_k^2} + \right.$$

$$\left. + i \left[\arctan \frac{l-a_k}{b_k} + \arctan \frac{l+a_k}{b_k} \right] \right\}$$

where $A_k = \frac{x_k^{2m}}{2nx_k^{2n-1}} = -\frac{1}{2n} x_k^{2m+1}$ since $x_k^{2n} = -1$. Passing to the limit for $l \to +\infty$ we obtain

$$\int_{-\infty}^{+\infty} \frac{x^{2m}}{1+x^{2n}}\, dx = \sum_{k=0}^{2n-1} \pm \pi i A_k$$

where the plus sign corresponds to $b_k > 0$ and the minus sign to $b_k < 0$. The integrals

$$A_k \int_{-\infty}^{+\infty} \frac{dx}{x-x_k} = A_k \left\{ \int_{-\infty}^{+\infty} \frac{(x-a_k)\, dx}{(x-a_k)^2+b_k^2} + i \int_{-\infty}^{+\infty} \frac{b_k\, dx}{(x-a_k)^2+b_k^2} \right\}, \quad k = 0, 1, \ldots$$

are obviously divergent and the numbers $\pm \pi i A_k = \lim_{l \to +\infty} \int_{-l}^{l} \frac{dx}{x-x_k}$ are their principal values.

Now note that $b_k > 0$ for $k = 0, 1, \ldots, n-1$ and $b_k < 0$ for $k = n, n+1, \ldots, 2n-1$. Hence, we can write

$$\int_{-\infty}^{+\infty} \frac{x^{2m}}{1+x^{2n}}\, dx = \pi i \left\{ \sum_{k=0}^{n-1} A_k - \sum_{k=n}^{2n-1} A_k \right\} \tag{A}$$

where

$$\sum_{k=0}^{n-1} A_k = -\frac{1}{2n} \sum_{k=0}^{n-1} x_k^{2m-1} = -\frac{1}{2n} \sum_{k=0}^{n-1} e^{i\frac{(2m+1)(2k+1)}{2n}\pi} =$$

$$= -\frac{1}{2n} \frac{e^{i\frac{2m+1}{2n}\pi} - e^{i\frac{(2m+1)(2n+1)}{2n}\pi}}{1-e^{i2\frac{2m+1}{2n}\pi}} =$$

$$= -\frac{1}{n} \frac{e^{i\frac{2m+1}{2n}\pi}}{1-e^{i2\frac{2m+1}{2n}\pi}} \tag{B}$$

because $e^{i(2m+1)\pi}=-1$. Next, putting $k=k'+n$ we obtain

$$\sum_{k=n}^{2n-1} A_k = -\frac{1}{2n}\sum_{k=n}^{2n-1} x_k^{2m+1} = -\frac{1}{2n}\sum_{k=n}^{2n-1} e^{i\frac{(2m+1)(2k+1)}{2n}\pi} =$$

$$= -\frac{1}{2n}\sum_{k'=0}^{n-1} e^{i\frac{(2m+1)(2k'+1)}{2n}\pi} e^{i(2m+1)\pi} =$$

$$= \frac{1}{2n}\sum_{k'=0}^{n-1} e^{i\frac{(2m+1)(2k'+1)}{2n}\pi} \qquad \text{(C)}$$

and the last sum differs from (B) only in its sign.

From (A), taking advantage of (B) and (C), we deduce

$$2J = \int_{-\infty}^{+\infty} \frac{x^{2m}}{1+x^{2n}}dx = -\frac{2\pi i}{n}\frac{e^{i\frac{2m+1}{2n}\pi}}{1-e^{i2\frac{2m+1}{2n}\pi}} = \frac{\pi}{n}\frac{1}{\sin\frac{2m+1}{2n}\pi}$$

Thus, the integrand $f(x)=\frac{x^{2m}}{1+x^{2n}}$ being an even function, we have (see Example 2)

$$J = \int_{0}^{+\infty} \frac{x^{2m}}{1+x^{2n}}\,dx = \frac{\pi}{2n}\frac{1}{\sin\frac{2m+1}{2n}\pi} \qquad (9.27')$$

4. The improper integral

$$\int_a^b \frac{dx}{x-c} = \lim_{\substack{\lambda\to 0+0\\ \mu\to 0+0}}\left\{\int_a^{c-\lambda}\frac{dx}{x-c}+\int_{c+\mu}^{b}\frac{dx}{x-c}\right\} = \ln\frac{b-c}{c-a}+\lim_{\substack{\lambda\to 0+0\\ \mu\to 0+0}}\ln\frac{\lambda}{\mu}$$

where $a<c<b$ is divergent. But if we put $\lambda=\mu>0$ and pass to the limit as $\lambda\to 0+0$ we see that this integral possesses the principal value

$$\text{v.p.}\int_a^b \frac{dx}{x-c} = \ln\frac{b-c}{c-a} \quad (a<c<b)$$

§ 4. IMPROPER MULTIPLE INTEGRALS

We shall first consider the case when the integrand is unbounded and the domain of integration is finite (bounded) and then pass to the case of an infinite (unbounded) domain of integration. For

simplicity's sake, we shall take double integrals although triple integrals and N-fold multiple integrals are considered similarly.

1. Integral of an Unbounded Function Over a Finite Domain. Let a function $f(M) = f(x, y)$ be defined over a finite domain Ω of the x, y-plane. We shall suppose that $f(x, y)$ is unbounded in every neighbourhood of a point $M_0(x_0, y_0) \in \bar{\Omega}$ and that for any

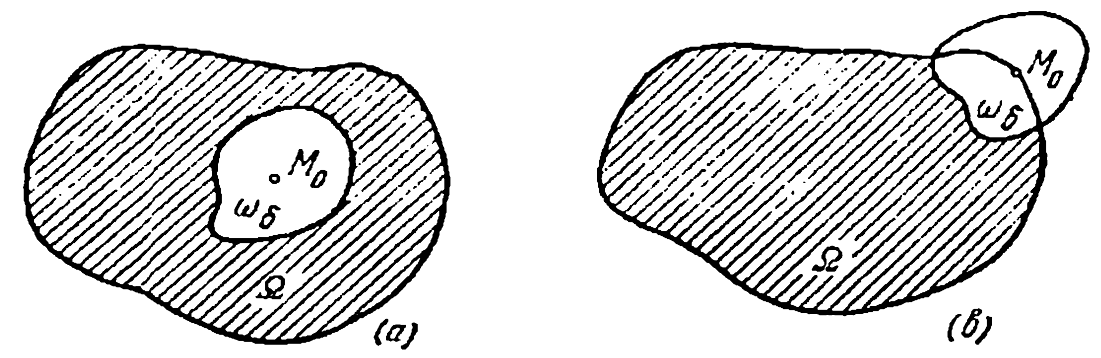


Fig. 9.4

domain $\Omega - \omega_\delta$ containing the point M_0 in its interior the function $f(x, y)$ is bounded and integrable in the ordinary sense over the domain $\Omega - \omega_\delta$ (shaded in Fig. 9.4a and b). This means that the integral $\iint\limits_{\Omega - \omega_\delta} f(M)\, d\omega$ is the limit of the corresponding integral sums (according to Definition 1 in § 2 of Chapter 1).* The subscript δ denotes the positive diameter of the domain ω_δ. If $\delta \to 0$ the domain ω_δ is contracted toward the point M_0.

Definition 1. The improper integral of the function $f(M) = f(x, y)$ *over the domain* Ω *is equal to the limit*

$$\lim_{\delta \to 0} \iint\limits_{\Omega - \omega_\delta} f(M)\, d\omega = \iint\limits_{\Omega} f(M)\, d\omega \tag{9.28}$$

If this limit exists, is finite and does not depend on the way the domain ω_δ *is contracted toward the point* M_0, *improper integral* (9.28) *is said to be* ***convergent***. *If otherwise, it is called* ***divergent***.

We say that the integral $\iint\limits_{\Omega - \omega_\delta} f(M)\, d\omega$ tends to a finite limit J, as $\delta \to 0$, which is independent of the way the domain ω_δ is con-

---

* The domains Ω and ω_δ and all the other domains which are considered in § 4 are supposed to be squarable. The symbol $\bar{\Omega}$ denotes the closure of Ω, i.e. a closed set which is the union of Ω and its boundary. The point M_0 may belong to the boundary of Ω or to its interior but it must be an interior point of ω_δ. The symbol $\Omega - \omega_\delta$ designates the set of all points belonging to Ω and not belonging to ω_δ. If Ω and ω_δ are squarable, the domain $\Omega - \omega_\delta$ is also squarable.

tracted to the point M_0 if for every sequence of domains

$$\omega_{\delta_1}, \omega_{\delta_2}, \ldots, \omega_{\delta_n}, \ldots \tag{9.29}$$

each of which contains the point M_0 in its interior and whose diameters satisfy the condition

$$\delta_n \to 0 \quad \text{for} \quad n \to +\infty^* \tag{9.30}$$

the corresponding number sequence

$$\iint\limits_{\Omega-\omega_{\delta_1}} f(M)\,d\omega, \quad \iint\limits_{\Omega-\omega_{\delta_2}} f(M)\,d\omega, \ \ldots, \ \iint\limits_{\Omega-\omega_{\delta_n}} f(M)\,d\omega, \ \ldots \tag{9.31}$$

converges to one and the same limit J irrespective of the choice of sequence (9.29).

Note 1. For an integral taken over a line segment $[a, b]$ (i.e. in the case $N = 1$) we took the intervals of the form $[a, b - \lambda]$ (which are connected point sets) as the domains $\Omega - \omega_{\delta_n}$. But if $N \geqslant 2$ the domains $\Omega - \omega_{\delta_n}$ and ω_{δ_n} are not necessarily supposed to be connected.

Definition 2. *Let the point M_0 lie in the interior or Ω. Suppose that integral (9.28) is divergent but sequence (9.31) tends to one and the same limit when (9.29) is an arbitrary sequence of concentric circles with centre at M_0 which are contracted toward M_0. Then this limit is called the **principal value of divergent integral** (9.28).***

Principal values of divergent improper double (and triple) integrals are applicable to some problems of mathematical physics.

Note 2. If M_0 is an interior point of Ω, then, when testing the integral $\iint\limits_{\Omega} f(M)\,d\omega$ for convergence, we can replace Ω by any subdomain $\Omega' \subset \Omega$ containing the point M_0 in its interior and consider the integral $\iint\limits_{\Omega'} f(M)\,d\omega$ instead of the original integral (compare this with the note after Definition 1 in § 1, Sec. 1). If the singular point M_0 belongs to the boundary of Ω we can take, as Ω', any subdomain which is the intersection of the domain Ω with an arbitrary domain Ω^* for which M_0 is an interior point.

---

* Here we do not suppose that (9.29) is a *monotone* sequence of sets, that is the one satisfying the condition $\omega_{\delta_1} \supset \omega_{\delta_2} \supset \ldots \supset \omega_{\delta_n} \supset \ldots$. The only requirement is that condition (9.30) must be fulfilled.

** Accordingly, the definition of the principal value of a divergent improper N-fold multiple integral involves the sequences of N-dimensional balls, instead of the sequences of circles which are contracted to the corresponding point.

Note 3. The case when $f(M)$ has an arbitrary finite number of singular points belonging to the domain Ω or to its boundary reduces to the case treated in Definition 1 if we appropriately break up the domain Ω into parts as it was done for onefold improper integrals.

2. Integrals of Nonnegative Functions. In this section we shall consider integrals of nonnegative functions which are simpler to investigate. The results obtained here will be further used for integrals of functions with alternating sign.

Theorem 9.5. Let the integrand $f(M) = f(x, y)$ *in integral* (9.28) *be a nonnegative function and let us take an arbitrary monotone sequence of concentric circles with centre at* M_0 *contracting toward* M_0 *as a sequence of type* (9.29). *Denoting a circle of radius* δ *with centre at* M_0 *by* K_δ *we can write this sequence in the form*

$$K_{\delta_1'} \supset K_{\delta_2'} \supset \ldots \supset K_{\delta_n'} \supset \ldots \text{ where } \delta_n' \to 0 \text{ for } n \to +\infty \quad (9.29')$$

Under these conditions, for integral (9.28) *to be convergent it is necessary and sufficient that the corresponding number sequence*

$$\iint\limits_{\Omega - K_{\delta_1'}} f(M)\, d\omega, \quad \iint\limits_{\Omega - K_{\delta_2'}} f(M)\, d\omega \quad \ldots, \quad \iint\limits_{\Omega - K_{\delta_n'}} f(M)\, d\omega, \ldots \quad (9.31')$$

be bounded.

Proof. Necessity directly follows from the definition of convergence of integral (9.28) because if integral (9.28) is convergent sequence (9.31′) must also be convergent, and hence it is bounded.

Sufficiency. Let sequence (9.31′) be bounded. Sequence (9.29′) being monotone and contracting to M_0, the sequence of domains of integration of integrals (9.31′) is monotone increasing, i.e. we have

$$\Omega - K_{\delta_1'} \subset \Omega - K_{\delta_2'} \subset \ldots \subset \Omega - K_{\delta_n'} \subset \ldots$$

Therefore number sequence (9.31′) is nondecreasing because the integrand $f(M) = f(x, y)$ is a nonnegative function. But this number sequence is supposed to be bounded and, consequently, it converges to a finite limit J:

$$\lim_{n \to +\infty} \iint\limits_{\Omega - K_{\delta_n'}} f(M)\, d\omega = J \quad (9.32)$$

and $\iint\limits_{\Omega - K_{\delta_n'}} f(M)\, d\omega \leqslant J$. To complete the proof of the theorem we must show that for any other choice of a sequence of domains (9.29) contracting toward M_0 the corresponding number sequence of form (9.31) converges to the same limit J. For every domain ω_{δ_n} entering into sequence (9.29) there exist circles $K_{\delta_r'}$ and $K_{\delta_q'}$ belonging

to sequence (9.29′) whose radii δ'_p and δ'_q tend to zero as $\delta_n \to 0$ such that

$$K_{\delta'_p} \supset \omega_{\delta_n} \supset K_{\delta'_q} \tag{9.33}$$

Relation (9.33) implies

$$\Omega - K_{\delta'_p} \subset \Omega - \omega_{\delta_n} \subset \Omega - K_{\delta'_q} \tag{9.34}$$

The integrand $f(M)$ being nonnegative, it follows from (9.34) that

$$\iint\limits_{\Omega - K_{\delta'_p}} f(M)\,d\omega \leqslant \iint\limits_{\Omega - \omega_{\delta_n}} f(M)\,d\omega \leqslant \iint\limits_{\Omega - K_{\delta'_q}} f(M)\,d\omega \tag{9.35}$$

But

$$\lim_{\delta'_p \to 0} \iint\limits_{\Omega - K_{\delta'_p}} f(M)\,d\omega = \lim_{\delta'_q \to 0} \iint\limits_{\Omega - K_{\delta'_q}} f(M)\,d\omega = J$$

and, consequently, relation (9.35) implies that

$$\lim_{n \to +\infty} \iint\limits_{\Omega - \omega_{\delta_n}} f(M)\,d\omega = J$$

which is what we set out to prove.

The following more general theorem is a direct consequence of Theorem 9.5.

Theorem 9.6. Let the integrand $f(M) = f(x, y)$ in integral (9.28) be a nonnegative function and let (9.29) be an arbitrary sequence of domains contracting to M_0 (see footnote on page 413). *Then integral (9.28) converges if and only if the corresponding numerical sequence of form (9.31) is bounded.*

Proof. Necessity is proved as in the foregoing theorem. To prove *sufficiency* let us take a monotone sequence of circles (9.29′) contracting toward M_0 and show that number sequence (9.31′) corresponding to this sequence of circles is bounded provided that sequence (9.31) is bounded. Then Theorem 9.5 will imply that integral (9.28) is convergent.

The fact that number sequence (9.31′) is bounded is proved a follows. Suppose that

$$\iint\limits_{\Omega - \omega_{\delta_m}} f(M)\,d\omega \leqslant C = \text{const} < +\infty \tag{9.36}$$

for all $m = 1, 2, 3, \ldots$. Since $\delta_m \to 0$ for $m \to +\infty$ we can assert that for any n there is m such that relation

$$K_{\delta'_n} \supset \omega_{\delta_m} \tag{9.37}$$

holds. Hence it follows that

$$\Omega - K_{\delta'_n} \subset \Omega - \omega_{\delta_n} \tag{9.38}$$

Therefore we obtain the inequality

$$\iint\limits_{\Omega - K_{\delta'_n}} f(M)\, d\omega \leqslant \iint\limits_{\Omega - \omega_{\delta_m}} f(M)\, d\omega \tag{9.39}$$

because $f(M) = f(x, y)$ is nonnegative. From (9.36) and (9.39) we conclude that the inequality

$$\iint\limits_{\Omega - K_{\delta'_n}} f(M)\, d\omega \leqslant C = \text{const} < +\infty \tag{9.40}$$

is fulfilled for all n which is what we set out to prove.

Example. Let us prove that the integral

$$\iint\limits_{\Omega} \frac{C}{r^{\alpha}}\, dx\, dy \text{ where } C = \text{const} > 0 \text{ and } r = \sqrt{(x - x_0)^2 + (y - y_0)^2} \tag{9.41}$$

taken over a finite domain Ω containing $M_0 \equiv (x_0, y_0)$ as its interior point converges for $\alpha < 2$ and diverges for $\alpha \geqslant 2$.

According to Note 2 in Sec. 1, integral (9.41) over the domain Ω can be replaced by the integral $\iint\limits_{\Omega'} \frac{C}{r^{\alpha}}\, dx\, dy$ taken over an arbitrary subdomain $\Omega' \subset \Omega$ for which M_0 is an interior point. Let us take a circle K_R of a sufficiently small radius R with centre at the point M_0 as a subdomain Ω'. Thus, we must investigate the integral

$$\iint\limits_{K_R} \frac{C}{r^{\alpha}}\, dx\, dy, \quad C > 0, \quad r = \sqrt{(x - x_0)^2 + (y - y_0)^2}, \quad \alpha = \text{const} \tag{9.42}$$

To do this let us choose a monotone sequence of circles

$$K_R \supset K_{\delta_1} \supset \ldots \supset K_{\delta_n} \supset \ldots \ni M_0 \text{ where } \delta_n \to 0 \text{ for } n \to +\infty \tag{9.43}$$

which contracts to M_0 and consider the integral

$$\iint\limits_{K_R - K_{\delta_n}} \frac{C}{r^{\alpha}}\, dx\, dy \tag{9.44}$$

Transforming this integral to polar coordinates we obtain

$$\iint\limits_{K_R - K_{\delta_n}} \frac{C}{r^{\alpha}}\, dx\, dy = \iint\limits_{K_R - K_{\delta_n}} \frac{C}{r^{\alpha}}\, r\, dr\, d\varphi = \int_0^{2\pi} d\varphi \int_{\delta_n}^{R} \frac{C}{r^{\alpha}}\, r\, dr =$$

$$= 2\pi C \int_{\delta_n}^{R} r^{1-\alpha}\, dr = \begin{cases} 2\pi C \left[\dfrac{r^{2-\alpha}}{2-\alpha} \right]_{r=\delta_n}^{r=R} & \text{for } \alpha \neq 2 \\ 2\pi C \, [\ln r]_{r=\delta_n}^{r=R} & \text{for } \alpha = 2 \end{cases} \tag{9.45}$$

Now, passing to the limit in (9.45) as $\delta_n \to 0$ we see that integral (9.40) is bounded for $\alpha < 2$ and approaches infinity for $\alpha \geqslant 2$. Consequently, integral (9.42) is convergent for $\alpha < 2$ and divergent for $\alpha \geqslant 2$, and hence the same conclusion applies to integral (9.41).

Similarly, in the case of N independent variables $x_1, x_2, \ldots, x_N$ the N-fold multiple integral

$$\iint_\Omega \ldots \int \frac{C}{r^\alpha} dx_1 \ldots dx_N, \quad C = \text{const} > 0, \qquad r = \sqrt{(x_1 - x_1^0)^2 + \ldots + (x_N - x_N^0)^2} \tag{9.46}$$

converges for $\alpha < N$ and diverges for $\alpha \geqslant N$ if $M_0 \equiv (x_1^0, \ldots, x_N^0)$ is an interior point of the N-dimensional domain Ω. Thus, the value $\alpha = N$ (equal to the dimension of the space) is a *critical* one in the sense that it separates the values of α ($\alpha < N$) for which integral (9.46) converges from the values of α ($\alpha \geqslant N$) for which it diverges, the value $\alpha = N$ corresponding to a divergent integral.

3. Absolute Convergence. Let a point M_0 belonging to a domain Ω be the only singular point of a function $f(M)$ defined in Ω. The point M_0 may be interior or belong to the boundary of the domain Ω. We suppose that for every domain ω for which M_0 is an interior point the function $f(M)$ is integrable in the ordinary sense over the domain $\Omega - \omega$.

Definition 3. *The integral* $\iint_\Omega f(M)\, d\omega$ *is said to be absolutely convergent if the integral* $\iint_\Omega |f(M)|\, d\omega$ *converges.*

Theorem 9.7. *If integral* $\iint_\Omega f(M)\, d\omega$ *converges absolutely it is convergent.*

Before proceeding to prove Theorem 9.7 we shall indicate some general properties of convergent improper integrals. We know that the limit of a sum is equal to the sum of the corresponding limits and that a constant factor can be taken outside the limit sign. Therefore we can assert that

(1) *if integrals* $\iint_\Omega f_1(M)\, d\omega$ *and* $\iint_\Omega f_2(M)\, d\omega$ *are convergent the integrals* $\iint_\Omega [f_1(M) \pm f_2(M)]\, d\omega$ *are also convergent and the equality*

$$\iint_\Omega [f_1(M) \pm f_2(M)]\, d\omega = \iint_\Omega f_1(M)\, d\omega \pm \iint_\Omega f_2(M)\, d\omega$$

holds;

(2) *if an integral* $\iint\limits_{\Omega} f(M)\,d\omega$ *converges, the integral* $\iint\limits_{\Omega} Cf(M)\,d\omega$ *where* $C=\text{const}$ *also converges and*

$$\iint\limits_{\Omega} Cf(M)\,d\omega = C\iint\limits_{\Omega} f(M)\,d\omega$$

We can now prove Theorem 9.7. Let us represent the integrand $f(M)$ as a difference of two nonnegative functions $f_1(M)=|f(M)|$ and $f_2(M)=|f(M)|-f(M)$:

$$f(M)=|f(M)|-[|f(M)|-f(M)]=f_1(M)-f_2(M) \quad (9.47)$$

By the hypothesis, the integral $\iint\limits_{\Omega} f_1(M)\,d\omega=\iint\limits_{\Omega}|f(M)|\,d\omega$ is convergent. We have

$$f_2(M)=|f(M)|-f(M)\leqslant 2\,|f(M)|$$

and, according to the conditions of the theorem, the integral

$$\iint\limits_{\Omega} 2\,|f(M)|\,d\omega = 2\iint\limits_{\Omega}|f(M)|\,d\omega$$

converges. Therefore, by Theorem 9.6, for every contracting sequence (9.29) the corresponding sequence of integrals $\iint\limits_{\Omega-\omega_{\delta_n}} 2\,|f(M)|\,d\omega$ is bounded. Furthermore, we have an obvious inequality

$$\iint\limits_{\Omega-\omega_{\delta_n}} f_2(M)\,d\omega \leqslant \iint\limits_{\Omega-\omega_{\delta_n}} 2\,|f(M)|\,d\omega$$

and hence the sequence of integrals $\iint\limits_{\Omega-\omega_{\delta_n}} f_2(M)\,d\omega$ is also bounded. Thus, by Theorem 9.6, the integral $\iint\limits_{\Omega} f_2(M)\,d\omega$ is convergent. But then relation (9.47) implies that the integral $\iint\limits_{\Omega} f(M)\,d\omega$ is also convergent and the equality

$$\iint\limits_{\Omega} f(M)\,d\omega=\iint\limits_{\Omega} f_1(M)\,d\omega-\iint\limits_{\Omega} f_2(M)\,d\omega \quad (9.48)$$

is fulfilled. The theorem has thus been proved.

Note. In the case of an improper N-fold multiple integral, for $N\geqslant 2$, the converse of the above theorem is also true, that is con-

vergence and absolute convergence are equivalent in the case $N \geqslant 2$ (see Sec. 5).

4. Tests for Absolute Convergence.

Theorem 9.8 (General Comparison Test). Let the inequality

$$0 \leqslant |f(M)| \leqslant g(M) \tag{9.49}$$

hold everywhere in the domain Ω. Besides, suppose that M_0 is the only singular point of the functions $f(M)$ and $g(M)$ in the domain Ω which may belong to its boundary or be an interior point. Then

(1) *if the integral $\iint\limits_{\Omega} g(M)\,d\omega$ converges the integral $\iint\limits_{\Omega} f(M)\,d\omega$ converges absolutely;*

(2) *if the integral $\iint\limits_{\Omega} f(M)\,d\omega$ diverges the integral $\iint\limits_{\Omega} g(M)\,d\omega$ also diverges.*

Proof. Let us take an arbitrary sequence of domains (9.29) contracting toward M_0. From inequality (9.49) it follows that

$$\iint\limits_{\Omega-\omega_{\delta_n}} |f(M)|\,d\omega \leqslant \iint\limits_{\Omega-\omega_{\delta_n}} g(M)\,d\omega \tag{9.50}$$

(1) If the integral $\iint\limits_{\Omega} g(M)\,d\omega$ is convergent the sequence $\left\{\iint\limits_{\Omega-\omega_{\delta_n}} |g(M)|\,d\omega\right\}$ is bounded and then, by inequality (9.50), the sequence $\left\{\iint\limits_{\Omega-\omega_{\delta_n}} |f(M)|\,d\omega\right\}$ is also bounded. Consequently, by Theorem 9.6, the integral $\iint\limits_{\Omega} |f(M)|\,d\omega$ is convergent.

(2) If the integral $\iint\limits_{\Omega} f(M)\,d\omega$ diverges the integral $\iint\limits_{\Omega} |f(M)|\,d\omega$ also diverges. Indeed, if the latter were convergent, the integral $\iint\limits_{\Omega} f(M)\,d\omega$ would also be convergent (by Theorem 9.7). The integral $\iint\limits_{\Omega} |f(M)|\,d\omega$ diverging, Theorem 9.6 implies that for any choice of a sequence of domains (9.29) contracting to M_0 the sequence $\iint\limits_{\Omega-\omega_{\delta_n}} |f(M)|\,d\omega$ is unbounded. But then, by inequality (9.50), the sequence $\iint\limits_{\Omega-\omega_{\delta_n}} g(M)\,d\omega$ is also unbounded, and conse-

quently the integral $\iint\limits_{\Omega} g(M)\,d\omega$ diverges, which is what we set out to prove.

Theorem 9.9 (Special Comparison Test). If a function $f(M)$ defined in Ω and having only one singular point $M_0(x_0, y_0)$ belonging to the interior of Ω or to its boundary satisfies the inequality

$$
\left.\begin{array}{c} |f(M)| = |f(x, y)| < \dfrac{C}{r^{\alpha}}, \quad C = \text{const} > 0, \\ r = \sqrt{(x-x_0)^2 + (y-y_0)^2} \end{array}\right. \tag{9.51}
$$

where $\alpha < 2$, the integral $\iint\limits_{\Omega} f(M)\,d\omega$ is absolutely convergent.

Proof. Integral (9.41) being convergent for $\alpha < 2$, we conclude, by Theorem 9.8 and inequality (9.51), that the integral $\iint\limits_{\Omega} |f(M)|\,d\omega$ is also convergent, which is what we set out to prove.

Note. In the case of an improper N-fold multiple integral

$$
\iint\limits_{\Omega}\cdots\int f(x_1, \ldots, x_N)\,dx_1\ldots dx_N
$$

taken over an N-dimensional domain Ω where the function $f(M) = f(x_1, \ldots, x_N)$ has only one singular point $M_0 = (x_1^0, \ldots, x_N^0)$ in the interior of the domain Ω or on its boundary we should take, in the special comparison test (Theorem 9.9), $r = \sqrt{(x_1-x_1^0)^2 + \ldots + (x_N-x_N^0)^2}$ and $\alpha < N$.

Example. Let us compute the force with which a material point $M_0(x_0, y_0, z_0)$ of unit mass is attracted by a material body, occupying a domain Ω in the x, y, z-space, with volume mass density $\rho(M) = \rho(x, y, z)$.

Let us write down the expressions of the projections on the coordinate axes of the force of attraction (see Sec. 5 in § 2 of Chapter 2):

$$
\left.\begin{aligned} F_x &= \iiint\limits_{\Omega} \rho(M)\,\frac{x-x_0}{r^3}\,dx\,dy\,dz \\ F_y &= \iiint\limits_{\Omega} \rho(M)\,\frac{y-y_0}{r^3}\,dx\,dy\,dz \\ F_z &= \iiint\limits_{\Omega} \rho(M)\,\frac{z-z_0}{r^3}\,dx\,dy\,dz \end{aligned}\right\} \tag{9.52}
$$

where

$$
r = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}, \quad M \equiv (x, y, z)
$$

In Chapter 2 we limited ourselves to the case when the point $M_0 \equiv (x_0, y_0, z_0)$ lies outside the body Ω. If M_0 belongs to the boundary of Ω or is its interior point the integrals (9.52) are improper in the general case. Suppose that the density $\rho(M) = \rho(x, y, z)$ is bounded on Ω, that is $0 \leqslant \rho(M) \leqslant \rho_0 =$ const for all the points $M \in \Omega$. Then

$$\left|\rho(M)\frac{x-x_0}{r^3}\right| \leqslant \rho_0 \left|\frac{1}{r^2}\,\frac{x-x_0}{r}\right| \leqslant \frac{\rho_0}{r^2} \text{ since } \left|\frac{x-x_0}{r}\right| \leqslant 1$$

Here we have $\alpha = 2 < N = 3$ and therefore, by the special comparison test, the first integral (9.52) is absolutely convergent. The other two integrals (9.52) are also absolutely convergent, which is proved similarly.

In the case $N \geqslant 2$ the improper N-fold multiple integrals possess a remarkable property which does not extend to the case $N = 1$. Namely, if $N \geqslant 2$, ordinary convergence of an improper integral implies its absolute convergence, that is the converse of Theorem 9.7 is true in this case.

5. Equivalence of Convergence and Absolute Convergence in the Case of Improper Multiple Integral. An improper integral of a function $f(M)$ converges for $N \geqslant 2$ if and only if the integral of $|f(M)|$ converges. This follows from Theorem 9.7 and the following theorem.

Theorem 9.10. If $N \geqslant 2$ and the integral $\overbrace{\int\int\ldots\int_{\Omega}}^{N} f(M)\,dx_1\ldots dx_N$ *converges, the integral* $\overbrace{\int\int\ldots\int_{\Omega}}^{N} |f(M)|\,dx_1\ldots dx_N$ *also converges.*

Proof. To simplify the notation we shall take the case $N = 2$ although the argument below is valid for any $N \geqslant 2$. Let a singular point M_0 of a function $f(M) = f(x, y)$ be an interior point of a plane region Ω the function is defined in.* Suppose the integral $\iint_{\Omega} f(M)\,d\omega$ is convergent while the integral $\iint_{\Omega} |f(M)|\,d\omega$ is divergent. Then we can take an arbitrary sequence $\{K_n\}$ of concentric circles

$$\Omega \supset K_1 \supset K_2 \supset \ldots \supset K_n \supset \ldots \ni M_0 \qquad (9.53)$$

* Without loss of generality (see Note 3 in § 4, Sec. 1), we can suppose that M_0 is the only singular point of $f(M)$ in Ω. If M_0 lies on the boundary of Ω we simply take, instead of K_n in (9.53), the intersections of the circles K_n with the domain Ω, i.e. their parts belonging to Ω.

with centre at M_0 contracting toward M_0 and write the relation

$$\lim_{n\to+\infty}\iint\limits_{\Omega-K_n}|f(M)|\,d\omega=+\infty \tag{9.54}$$

because $|f(M)|$ is nonnegative. Sequence (9.53) can therefore be chosen in such a way that the inequalities

$$\iint\limits_{K_n-K_{n+1}}|f(M)|\,d\omega\geqslant 2\iint\limits_{\Omega-K_n}|f(M)|\,d\omega+2n,\quad n=1,\ 2,\ \ldots \tag{9.55}$$

hold. Let us introduce the functions

$$f_+(M)=\frac{1}{2}[|f(M)|+f(M)]\quad\text{and}\quad f_-(M)=\frac{1}{2}[|f(M)|-f(M)] \tag{9.56}$$

We obviously have $f_+(M)\geqslant 0$, $f_-(M)\geqslant 0$ and

$$f(M)=f_+(M)-f_-(M),\quad |f(M)|=f_+(M)+f_-(M) \tag{9.57}$$

From (9.57) it follows that

$$\iint\limits_{K_n-K_{n+1}}|f(M)|\,d\omega=\iint\limits_{K_n-K_{n+1}}f_+(M)\,d\omega+\iint\limits_{K_n-K_{n+1}}f_-(M)\,d\omega \tag{9.58}$$

We shall suppose that sequence (9.53) is chosen in such a way that

$$\iint\limits_{K_n-K_{n+1}}f_+(M)\,d\omega\geqslant\iint\limits_{K_n-K_{n+1}}f_-(M)\,d\omega \tag{9.59}$$

(if otherwise we can pass from sequence (9.53) to an appropriate subsequence and replace, if necessary, $f(M)$ by $-f(M)$). Then relations (9.55) and (9.58) imply that

$$\iint\limits_{K_n-K_{n+1}}f_+(M)\,d\omega>\iint\limits_{\Omega-K_n}|f(M)|\,d\omega+n,\quad n=1,\ 2,\ \ldots \tag{9.60}$$

If we break up the annulus K_n-K_{n+1} into sufficiently small squarable cells $\Delta\omega_i$ the lower Darboux sum $\sum m_i^{f_+}\Delta\omega_i^*$ corresponding to the integral $\iint\limits_{K_n-K_{n+1}}f_+(M)\,d\omega$ satisfies, by (9.60), the inequality

$$\sum m_i^{f_+}\Delta\omega_i>\iint\limits_{\Omega-K_n}|f(M)|\,d\omega+n,\quad n=1,\ 2,\ \ldots \tag{9.61}$$

---

* Here $m_i^{f_+}$ designates the greatest lower bound of $f_+(M)$ on the cell $\Delta\omega_i\subset K_n-K_{n+1}$.

We have $m_i^{f+} \geqslant 0$ for all these cells since $f_+ \geqslant 0$ everywhere in Ω. Let us delete all the terms in the sum $\sum m_i^{f+} \Delta\omega_i$ for which $m_i^{f+} = 0$ (this of course does not affect the validity of inequality (9.61)). Denoting by G_n the domain which is the union of the remaining cells, we can write $f(M) = f_+(M)$ for $M \in G_n$, and

$$\iint_{G_n} f(M)\, d\omega = \iint_{G_n} f_+(M)\, d\omega \geqslant \sum_{\Delta\omega_i \subset G_n} m_i^{f+}\Delta\omega_i > \iint_{\Omega - K_n} |f(M)|\, d\omega + n,$$
$$n = 1, 2, \ldots \tag{9.62}$$

Furthermore, we have

$$\iint_{\Omega - K_n} f(M)\, d\omega \geqslant -\iint_{\Omega - K_n} |f(M)|\, d\omega, \quad n = 1, 2, \ldots \tag{9.63}$$

Adding up (9.62) and (9.63) we derive

$$\iint_{H_n} f(M)\, d\omega > n, \qquad n = 1, 2, \ldots \tag{9.64}$$

where $H_n = (\Omega - K_n) + G_n$. If we denote by ω_n the difference $\Omega - H_n$, the diameter of ω_n tends to zero for $n \to +\infty$. Consequently, (9.64) implies that the integral $\iint\limits_{\Omega} f(M)\, d\omega$ diverges, which contradicts the hypothesis. Thus, supposing that the integral $\iint\limits_{\Omega} |f(M)|\, d\omega$ is divergent we arrive at a contradiction and hence it is convergent. The theorem has thus been proved.

Note. If, in the case $N \geqslant 2$, we consider the domains $\Omega - \omega_{\delta_n}$ entering into the definition of an improper N-fold multiple integral to be connected Theorem 9.10 remains true. In fact, the domain $H_n = (\Omega - K_n) + G_n$ $(n = 1, 2, \ldots)$ appearing in the above proof of Theorem 9.10 can be made connected without violating the validity of inequality (9.64). For this purpose it is sufficient to join together the connected components constituting H_n by squarable strips of sufficiently small total area. The possibility of constructing such strips becomes obvious if we break up the annulus $K_n - K_{n+1}$ into squarable parts entering into the integral sum $\sum\limits_{\Delta\omega_i \subset K_n - K_{n+1}} m_i^{f+}\Delta\omega_i$ by means of rays starting from the centre M_0 of the annulus and concentric circles with centre at M_0.

In contrast to this, if in the case $N = 1$, i.e. for an improper integral $\int\limits_a^b f(x)\, dx$ taken over an interval $[a, b]$, we take, instead of the sequences of intervals of the form $[a, b - \lambda]$ $(\lambda > 0,\ \lambda <$

$< b - a$) entering into the corresponding definition (see relation (9.17) in § 2), exhaustive sequences of arbitrary disconnected domains, this will lead to a narrower class of functions for which the improper integrals exist. Indeed, in this case only the functions absolutely integrable in the improper sense will constitute the class of functions with the convergent improper integrals and thus the functions the integrals of which are conditionally convergent will be excluded (the classes of functions absolutely integrable in the improper sense will obviously be the same in both definitions).

6. Improper Integrals with Infinite Domain of Integration. The integrals over unbounded domains whose integrands are functions bounded in any finite domain are investigated in quite analogous fashion. As an example, we shall formulate the definition of an improper integral over an unbounded domain and a sufficient condition for convergence.

Definition 4. Let Ω be an infinite (unbounded) domain. A sequence of finite (bounded) subdomains

$$\Omega_1,\ \Omega_2,\ \ldots,\ \Omega_n,\ \ldots \tag{9.65}$$

*is said to be **exhaustive** if for any $R > 0$ there is $m = m(R)$ such that all the points of the domain Ω lying within the circle of radius R with centre at the origin belong to all Ω_n for $n > m(R)$.*

Definition 5. Let a function $f(M)$ defined in an infinite domain Ω be integrable in the ordinary sense on every finite subdomain. If for any choice of exhaustive sequence (9.65) *the corresponding number sequence*

$$\iint\limits_{\Omega_1} f(M)\,d\omega,\quad \iint\limits_{\Omega_2} f(M)\,d\omega,\ \ldots,\ \iint\limits_{\Omega_n} f(M)\,d\omega,\ \ldots$$

*converges to one and the same finite limit J the integral $\iint\limits_{\Omega} f(M)\,d\omega$ is said to be **convergent**: if otherwise the integral is called **divergent**.*

Sufficient Condition for Convergence. If $f(M) = f(x, y)$ satisfies the requirements of the foregoing definition and the inequality

$$|f(M)| \leqslant \frac{C}{r^\alpha} \text{ where } C = \text{const} > 0,$$

$$\alpha = \text{const} > 2 \text{ and } r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$$

is fulfilled where $M_0 = (x_0, y_0)$ is a fixed point belonging to Ω the integral $\iint\limits_{\Omega} f(M)\,d\omega$ is convergent.

In conclusion we note that the general theorems analogous to Theorems 9.5, 9.6, 9.7, 9.8 and 9.10 also apply to improper integrals with unbounded domains of integration.

7. Methods of Computing Improper Multiple Integrals. A convergent improper double integral can be reduced to the corresponding twofold iterated integral like a proper double integral under the following conditions:

(1) if the integrand is a *nonnegative* (*nonpositive*) function it is required that the iterated integral of this function be convergent;

(2) if the integrand is a *function with alternating sign* it is supposed that the iterated integral of its modulus converges.*

The method of changing variables is applied to a convergent improper N-fold multiple integral according to the same rules as in the case of a proper N-fold multiple integral.

Here we do not present the proofs of these general assertions and confine ourselves to an example in which we encounter reduction of an improper double integral to an iterated one and change of variables in an improper integral.

Let it be necessary to evaluate the Euler-Poisson integral $J = \int_0^{+\infty} e^{-x^2}\,dx$ (see also Example 1 in Sec. 4, § 1 where the convergence of this integral was established). The value of a definite integral remains the same when the notation of the variable of integration is changed and therefore $J = \int_0^{+\infty} e^{-x^2}\,dx = \int_0^{+\infty} e^{-y^2}\,dy$. Hence we can write

$$J^2 = \int_0^{+\infty} e^{-x^2}\,dx \int_0^{+\infty} e^{-y^2}\,dy = \int_0^{+\infty} \left(e^{-x^2} \int_0^{+\infty} e^{-y^2}\,dy\right) dx =$$

$$= \int_0^{+\infty} dx \int_0^{+\infty} e^{-(x^2+y^2)}\,dy$$

and the iterated integral is convergent. The double integral

$$\int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}\,dx\,dy$$

---

* The situation is similar in the case of an improper N-fold multiple integral for $N \geqslant 3$.

is also convergent which is implied by the sufficient condition for convergence given in Sec. 6. Consequently,

$$J^2 = \int_0^{+\infty} \int_0^{+\infty} e^{-(x^2+y^2)}\, dx\, dy$$

Passing to polar coordinates we obtain

$$J^2 = \int_0^{\frac{\pi}{2}} d\varphi \int_0^{+\infty} e^{-r^2} r\, dr = \frac{\pi}{2} \int_0^{+\infty} e^{-r^2} r\, dr = \frac{\pi}{4}$$

Hence, we have

$$J = \int_0^{+\infty} e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}$$

This technique of evaluating the integral was developed by Poisson.

10 Integrals Dependent on Parameter

In this chapter we consider the properties of integrals dependent on a parameter which are effectively used in analytical methods of mathematics and mathematical physics. Such important integrals as *Euler's integrals of the first and the second kind* (the *beta* and *gamma functions*; see § 3), integrals of the type of a potential function etc. belong to the class of integrals dependent on a parameter.

§ 1. PROPER AND SIMPLEST IMPROPER INTEGRALS DEPENDENT ON PARAMETER

1. Proper Integrals Dependent on Parameter. Let $u = f(x, y)$ be a function defined in a rectangle Π: $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$. We suppose that this function is integrable with respect to x on the interval $a \leqslant x \leqslant b$ for every value of y belonging to the interval $c \leqslant y \leqslant d$. Then the integral

$$J(y) = \int_a^b f(x, y)\,dx \qquad (10.1)$$

(dependent on the parameter y) is a function of the parameter y on the interval $c \leqslant y \leqslant d$. Here we shall study the properties of integrals (10.1).

Theorem 10.1 (On the Continuity of an Integral Dependent on a Parameter with Respect to the Parameter). *If the function $f(x, y)$ is continuous in the closed rectangle Π: $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$, the integral $J(y) = \int_a^b f(x, y)\,dx$ is a continuous function of the parameter y on the interval $c \leqslant y \leqslant d$.*

Proof. Since the function $f(x, y)$ is continuous in the closed rectangle Π it is uniformly continuous. Hence, for every $\varepsilon > 0$ there is $\delta = \delta(\varepsilon) > 0$ such that the inequalities

$$|x' - x''| < \delta(\varepsilon) \quad \text{and} \quad |y' - y''| < \delta(\varepsilon) \qquad (10.2)$$

imply the inequality

$$|f(x', y') - f(x'', y'')| < \frac{\varepsilon}{b-a} \tag{10.3}$$

(here $\delta(\varepsilon)$ depends solely on ε and is independent of the positions occupied by the points (x', y') and (x'', y'') within the rectangle Π provided that inequalities (10.2) are fulfilled). In particular, putting $x' = x'' = x$ we see that for any y' and y'' belonging to the interval $c \leqslant y \leqslant d$ and satisfying the inequality

$$|y' - y''| < \delta(\varepsilon) \tag{10.2'}$$

and for all x, $a \leqslant x \leqslant b$, the inequality

$$|f(x, y') - f(x, y'')| < \frac{\varepsilon}{b-a} \tag{10.3'}$$

holds. Therefore, for any y' and y'' belonging to the interval $c \leqslant y \leqslant d$ and satisfying inequality (10.2′) the inequality

$$|J(y') - J(y'')| = \left| \int_a^b [f(x, y') - f(x, y'')]\, dx \right| \leqslant$$

$$\leqslant \int_a^b |f(x, y') - f(x, y'')|\, dx \leqslant \frac{\varepsilon}{b-a}(b-a) = \varepsilon$$

is fulfilled, which means that $J(y)$ is *uniformly continuous* on the interval $c \leqslant y \leqslant d$. The theorem has been proved.

Corollary. Under the conditions of Theorem 10.1, *the function* $F(u, v, y) = \int_u^v f(x, y)\, dx$ *is continuous in the closed rectangle*

$$\Pi^*: a \leqslant u \leqslant b, \quad a \leqslant v \leqslant b, \quad c \leqslant y \leqslant d$$

Proof. The function $f(x, y)$ being continuous in the closed rectangle Π, there is a constant C, $0 < C < +\infty$, such that $|f(x, y)| < C$ everywhere in Π. Therefore, for any points (u', v', y') and (u'', v'', y'') of Π^* we have the following inequality:

$$|F(u', v', y') - F(u'', v'', y'')| = \left| \int_{u'}^{v'} f(x, y')\, dx - \int_{u''}^{v''} f(x, y'')\, dx \right| \leqslant$$

$$\leqslant \left| \int_{u'}^{v'} [f(x, y') - f(x, y'')]\, dx \right| + \left| \int_{u'}^{u''} f(x, y'')\, dx \right| +$$

$$+ \left| \int_{v'}^{v''} f(x, y'')\, dx \right| \leqslant \left| \int_{u'}^{v'} [f(x, y') - f(x, y'')]\, dx \right| +$$

$$+ C|u'' - u'| + C|v'' - v'| \tag{10.4}$$

Now let the point (u', v', y') be fixed and the point (u'', v'', y'') tend to (u', v', y'): $(u'', v'', y'') \to (u', v', y')$. Then, by Theorem 10.1 the first term on the right-hand side of (10.4) tends to zero. The second and the third terms on the right-hand side of (10.4) also tending to zero as $(u'', v'', y'') \to (u', v', y')$, the theorem has thus been proved.

Theorem 10.2 (On Differentiation of an Integral Dependent on a Parameter with Respect to the Parameter). *If the function $f(x, y)$ and its partial derivative $f'_y(x, y)$ are continuous in the rectangle* $\Pi: a \leqslant x \leqslant b,\ c \leqslant y \leqslant d$, *the integral $J(y) = \int\limits_a^b f(x, y)\,dx$ is a differentiable function of the parameter y on the interval $c \leqslant y \leqslant d$, and the relation*

$$\frac{dJ}{dy} = \frac{d}{dy}\int\limits_a^b f(x, y)\,dx = \int\limits_a^b f'_y(x, y)\,dx \tag{10.5}$$

is valid for all y belonging to this interval.

Note. Formula (10.5) is known as **Leibniz' rule (formula)** for differentiating an integral with respect to the parameter it depends on: *the derivative of the integral with respect to the parameter is equal to the integral of the derivative of the integrand with respect to this parameter.*

Proof. We must show that

$$\lim_{\Delta y \to 0} \frac{J(y+\Delta y) - J(y)}{\Delta y} = \int\limits_a^b f'_y(x, y)\,dx$$

For this purpose we shall prove that the difference between the variable quantity $\frac{J(y+\Delta y)-J(y)}{\Delta y}$ $(\Delta y \neq 0)$ and the integral $\int\limits_a^b f'_y(x, y)\,dx$ tends to zero when $\Delta y \to 0$. By virtue of Lagrange's formula of finite increments, we have

$$\frac{J(y+\Delta y)-J(y)}{\Delta y} = \int\limits_a^b \frac{f(x, y+\Delta y)-f(x, y)}{\Delta y}\,dx = \int\limits_a^b f'_y(x, y+\theta\Delta y)\,dx$$

where $0 < \theta < 1$. Consequently, we can write

$$\frac{J(y+\Delta y)-J(y)}{\Delta y} - \int\limits_a^b f'_y(x, y)\,dx = \int\limits_a^b [f'_y(x, y+\theta\Delta y) - f'_y(x, y)]\,dx \tag{10.6}$$

Let us estimate the above difference for sufficiently small values of $|\Delta y|$. Let there be given $\varepsilon > 0$. Since the derivative $f'_y(x, y)$ is continuous on the closed rectangle Π it is uniformly continuous on it and therefore there is $\delta(\varepsilon) > 0$ such that, for $|\Delta y| < \delta(\varepsilon)$, the inequality

$$|f'_y(x, y+\Delta y) - f'_y(x, y)| < \frac{\varepsilon}{b-a}$$

holds for all $x \in [a, b]$ and any y and $y + \Delta y$ belonging to the interval $[c, d]$. We have $0 < \theta < 1$ and hence for all x, y and $y + \Delta y$ mentioned above the inequality

$$|f'_y(x, y+\theta\Delta y) - f'_y(x, y)| < \frac{\varepsilon}{b-a}$$

is fulfilled. Thus, by (10.6), we have

$$\left|\frac{J(y+\Delta y) - J(y)}{\Delta y} - \int_a^b f'_y(x, y)\,dx\right| = \left|\int_a^b [f'_y(x, y+\theta\Delta y) - f'_y(x, y)]\,dx\right| \leqslant$$

$$\leqslant \int_a^b \left|f'_y(x, y+\theta\Delta y) - f'_y(x, y)\right| dx <$$

$$< \frac{\varepsilon}{b-a}(b-a) = \varepsilon$$

for all $|\Delta y| < \delta(\varepsilon)$. The theorem has been proved.

Theorem 10.3 (On Differentiating an Integral Dependent on a Parameter Whose Limits of Integration Also Depend on the Parameter with Respect to the Parameter). *Let $f(x, y)$ and $f'_y(x, y)$ be continuous in the rectangle Π: $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$, and let $x = x_1(y)$ and $x = x_2(y)$ be differentiable functions defined on the interval $c \leqslant y \leqslant d$ and satisfying the condition $a < x_i(y) < b$ $(i = 1, 2)$. Then the derivative of the integral*

$$J(y) = \int_{x_1(y)}^{x_2(y)} f(x, y)\,dx \tag{10.7}$$

with respect to the parameter y exists and is equal to

$$J'(y) = \int_{x_1(y)}^{x_2(y)} f'_y(x, y)\,dx + f(x_2(y), y)\frac{dx_2}{dy} - f(x_1(y), y)\frac{dx_1}{dy} \tag{10.8}$$

Proof. We have

$$J(y) = F(x_1(y), x_2(y), y) \tag{10.9}$$

where the function $F(u, v, y) \equiv \int\limits_u^v f(x, y)\,dx$ considered for $a \leqslant u \leqslant b$, $a \leqslant v \leqslant b$ and $c \leqslant y \leqslant d$ possesses the continuous partial derivatives

$$F'_u = -f(u, y), \quad F'_v = f(v, y), \quad F'_y = \int\limits_u^v f'_y(x, y)\,dx \quad (10.10)$$

By Theorem 10.3, the partial derivative F'_y exists, and the Corollary of Theorem (10.1) implies that it is continuous. The functions $x = x_1(y)$ and $x = x_2(y)$ being differentiable, we can apply the rule for differentiating a composite function to integral (10.8) which results in equality (10.8). The theorem has been proved.

Theorem 10.4 (On Integration of an Integral Dependent on a Parameter with Respect to the Parameter). *If $f(x, y)$ is a continuous function in the rectangle* Π: $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$, *we have*

$$\int\limits_c^d J(y)\,dy = \int\limits_c^d dy \int\limits_a^b f(x, y)\,dx = \int\limits_a^b dx \int\limits_c^d f(x, y)\,dy \qquad (10.11)$$

which means that to integrate the integral $J(y) = \int\limits_a^b f(x, y)\,dx$ *with respect to the parameter y we can integrate the integrand $f(x, y)$ with respect to this parameter.*

Proof. Equality (10.11) is a consequence of the theorem on reducing a double integral to an iterated one (see Chapter 1, § 5).

We shall present here another proof of the theorem which can easily be extended to the case of an arbitrary dimension N (see § 4). Moreover, instead of equality (10.11) we shall establish a more general relation

$$\int\limits_c^d dy \int\limits_a^t f(x, y)\,dx = \int\limits_a^t dx \int\limits_c^d f(x, y)\,dy \quad \text{for } a \leqslant t \leqslant b \quad (10.12)$$

Introducing the notation

$$\varphi(t) = \int\limits_c^d dy \int\limits_a^t f(x, y)\,dx, \quad \psi(t) = \int\limits_a^t dx \int\limits_c^d f(x, y)\,dy \quad (10.13)$$

we see that it is sufficient to prove that $\varphi'(t) \equiv \psi'(t)$ for $a \leqslant t \leqslant b$ and that $\varphi(a) = \psi(a)$ because this obviously implies the identity $\varphi(t) = \psi(t)$, $t \in [a, b]$.

The equality $\varphi(a) = \psi(a)$ is apparent since $\varphi(a) = 0$ and $\psi(a) = 0$. Putting $F(t, y) = \int\limits_a^t f(x, y)\,dx$ we can write $\varphi(t) = \int\limits_c^d F(t, y)\,dy$ where the function $F(t, y)$ is continuous in the rectangle II^*: $a \leqslant t \leqslant b$, $c \leqslant y \leqslant d$, by virtue of the Corollary of Theorem 10.1. Furthermore, by the hypothesis of the theorem, the derivative $F_t'(t, y) = f(t, y)$ is continuous and, consequently, by Theorem 10.2, we have

$$\varphi'(t) = \int\limits_c^d F_t'(t, y)\,dy = \int\limits_c^d f(t, y)\,dy \tag{10.14}$$

Putting $\xi(x) = \int\limits_c^d f(x, y)\,dy$ we obtain $\psi(t) = \int\limits_a^t \xi(x)\,dx$. The function $\xi(x)$ being continuous in x on the interval $a \leqslant x \leqslant b$ (by Theorem 10.1), the theorem on differentiating a definite integral with respect to its upper limit of integration implies the relation

$$\psi'(t) = \frac{d}{dt}\int\limits_a^t \xi(x)\,dx = \xi(t) = \int\limits_c^d f(t, y)\,dy \tag{10.15}$$

From (10.14) and (10.15) it follows that $\varphi'(t) \equiv \psi'(t)$ for $a \leqslant t \leqslant b$. Hence, by the equality $\varphi(a) = \psi(a)$, we have $\varphi(t) \equiv \psi(t)$ for $a \leqslant t \leqslant b$. In particular, $\varphi(b) = \psi(b)$, that is equality (10.11) holds, which is what we set out to prove.

2. Simplest Improper Integrals Dependent on Parameter. Theorems 10.1, 10.2 and 10.4 can easily be extended to improper integrals of the following special type:

$$J(y) = \int\limits_a^b f(x, y)\,g(x)\,dx \tag{10.16}$$

where the function $f(x, y)$ is continuous and the function $g(x)$ may be discontinuous, in the general case, but such that the integral $\int\limits_a^b |g(x)|\,dx$ is convergent, including the case when one or both limits of integration are infinite.

Now we proceed to formulate the exact conditions of these generalized theorems*.

---

* Theorems 10.1′, 10.2′ and 10.4′ are used in mathematical physics and in the theory of Fourier's integral.

Theorem 10.1' (Generalized Theorem on the Continuous Dependence of an Integral on the Parameter). *If the function $f(x, y)$ is continuous and bounded for $a \leqslant x < +\infty$, $c \leqslant y \leqslant d$, and the integral $\int\limits_a^{+\infty} |g(x)|\,dx$ converges, the integral*

$$J(y) = \int\limits_a^{+\infty} f(x, y)\,g(x)\,dx \tag{10.17}$$

is a continuous function of y on the interval $c \leqslant y \leqslant d$.

Theorem 10.2' (Generalized Theorem on Differentiating an Integral with Respect to the Parameter). *If the function $f(x, y)$ and its partial derivative $f'_y(x, y)$ are continuous and bounded for $a \leqslant x < +\infty$, $c \leqslant y \leqslant d$, and the integral $\int\limits_a^{+\infty} |g(x)|\,dx$ is convergent, integral* (10.17) *is a differentiable function of the parameter y for $c \leqslant y \leqslant d$ and the equality*

$$J'(y) = \int\limits_a^{+\infty} f'_y(x, y)\,g(x)\,dx \tag{10.18}$$

holds for all $y \in [c, d]$.

Theorem 10.4' (Generalized Theorem on Integrating an Integral with Respect to the Parameter). *Under the conditions of Theorem* 10.1 *integral* (9.17) *is an integrable function of the parameter y on the interval $c \leqslant y \leqslant d$ and*

$$\int\limits_c^d J(y)\,dy = \int\limits_c^d dy \int\limits_a^{+\infty} f(x, y)\,g(x)\,dx =$$

$$= \int\limits_a^{+\infty} \left(g(x) \int\limits_c^d f(x, y)\,dy \right) dx \tag{10.19}$$

As an instance, let us prove Theorem 10.1'. Let $|f(x, y)| < C =$ $=\text{const}$ for $a \leqslant x < +\infty$, $c \leqslant y \leqslant d$ and $\int\limits_a^{+\infty} |g(x)|\,dx < K < +\infty$.

Let $\varepsilon > 0$ be taken arbitrarily. The integral $\int\limits_a^{+\infty} |g(x)|\,dx$ being convergent, there exists a sufficiently large $l > a$ such that $2C \int\limits_l^{+\infty} |g(x)|\,dx < \frac{\varepsilon}{2}$. Taking such a fixed value of l and choosing

arbitrary values y' and y'' belonging to the interval $c \leqslant y \leqslant d$ we can represent the difference $J(y') - J(y'')$ in the form

$$J(y')-J(y'')=\int_a^l [f(x, y')-f(x, y'')]\, g(x)\, dx+$$
$$+\int_l^{+\infty} [f(x, y')-f(x, y'')]\, g(x)\, dx \qquad (10.20)$$

Since the function $f(x, y)$ is continuous in the rectangle $a \leqslant x \leqslant l$, $c \leqslant y \leqslant d$, it is uniformly continuous. Therefore there exists $\delta = \delta(\varepsilon) > 0$ such that for any y' and y'' belonging to the interval $c \leqslant y \leqslant d$ and satisfying the inequality $|y' - y''| < \delta(\varepsilon)$ we have the inequality

$$|f(x, y')-f(x, y'')|<\frac{\varepsilon}{2K}$$

for all x, $a \leqslant x \leqslant l$. Then equality (10.20) implies the relation

$$|J(y')-J(y'')| \leqslant \int_a^l |f(x, y')-f(x, y'')|\, |g(x)|\, dx+$$
$$+\int_l^{+\infty} \{|f(x, y')|+|f(x, y'')|\}\, |g(x)|\, dx \leqslant$$
$$\leqslant \frac{\varepsilon}{2K}\int_a^l |g(x)|\, dx+2C\int_l^{+\infty} |g(x)|\, dx \leqslant$$
$$\leqslant \frac{\varepsilon}{2K} K+\frac{\varepsilon}{2}=\varepsilon \quad \text{for } |y'-y''|<\delta(\varepsilon)$$

which means that the integral $J(y)$ is a continuous function of y on the interval $c \leqslant y \leqslant d$.

Let the reader prove Theorems 10.2′ and 10.4′ for integrals (10.17) and also rephrase and prove Theorems 10.1′, 10.2′ and 10.4′ for integrals (10.16). We only note that in the case of integral (10.16) the boundedness of $f(x, y)$ and $f'_y(x, y)$ is implied by the continuity of $f(x, y)$ and $f'_y(x, y)$ in the domain $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$, and that it is unnecessary to break up the interval of integration $a \leqslant x \leqslant b$ into parts as it was done in the proof of Theorem 10.1′ for integrals of type (10.17).

Differentiation and integration of integrals with respect to parameters are widely applied to evaluating integrals dependent on parameters and also for computing integrals not involving parameters after a parameter is appropriately introduced.

Example. Let us evaluate the integral

$$J(y)=\int_0^{+\infty} e^{-\alpha x}\frac{\sin xy}{x}\,dx \quad \text{where } \alpha=\text{const}>0, \quad -A\leqslant y\leqslant A \tag{10.21}$$

Putting $f(x, y)=\frac{\sin xy}{x}$ and $g(x)=e^{-\alpha x}$ we see that $f(x, y)$ and $f'_y(x, y)$ are continuous and bounded for $0\leqslant x<+\infty$, $-A\leqslant y\leqslant A$, and the integral $\int_0^{+\infty}|g(x)|\,dx=\int_0^{+\infty}e^{-\alpha x}\,dx=\frac{1}{\alpha}$ is convergent. Therefore we can apply generalized Theorem 10.2′ for the integrals of form (10.17). Performing differentiation with respect to the parameter under the sign of integration we obtain

$$J'(y)=\int_0^{+\infty} e^{-\alpha x}\cos xy\,dx$$

Integrating by parts with respect to x in the latter integral twice we find

$$J'(y)=\frac{\alpha}{\alpha^2+y^2} \tag{10.22}$$

By (10.21), we have $J(0)=0$ and, therefore, integrating (10.22) from 0 to y we derive

$$J(y)=\int_0^y \frac{\alpha}{\alpha^2+y^2}\,dy=\arctan\frac{y}{\alpha}$$

There is another way of completing the evaluation of this integral. Namely, after relation (10.22) has been obtained, we can integrate it with respect to y which results in

$$J(y)=\arctan\frac{y}{\alpha}+C, \quad C=\text{const}$$

But, by (10.21), we have $J(0)=0$, and, consequently, putting $y=0$ in the above relation involving the constant C we see that $C=0$. Both approaches involve a known value of the integral in question for a particular value of the parameter y.

§ 2. IMPROPER INTEGRALS DEPENDENT ON PARAMETER

Let a function $u=f(x, y)$ be defined for $0\leqslant x<+\infty$, $c\leqslant y\leqslant d$ and let, for every value of $y\in[c, d]$, the integral

$$J(y)=\int_a^{+\infty} f(x, y)\,dx \tag{10.23}$$

be convergent. Then $J(y)$ is a function of y defined on the interval $[c, d]$. According to the definition of the improper integral, we have

$$J(y)=\int_a^{+\infty} f(x, y)\,dx=\lim_{l\to+\infty}\int_a^l f(x, y)\,dx \tag{10.24}$$

In the case of integrals of unbounded functions we have a similar situation. For instance, let a function $u=f(x, y)$ defined for $a \leqslant x<b$, $c \leqslant y \leqslant d$ be unbounded for $x \to b-0$ and let the integral

$$J^*(y)=\int_a^b f(x, y)\,dx=\lim_{\lambda\to 0+0}\int_a^{b-\lambda} f(x, y)\,dx \tag{10.25}$$

be convergent for every value of y belonging to the interval $[c, d]$. Then $J^*(y)$ is a function of y defined on the interval $[c, d]$.

1. Uniform Convergence. The notion of *uniform convergence* plays an important role in the theory of improper integrals dependent on a parameter. Uniformly convergent improper integrals can be operated on like proper ones (see Sec. 3 of § 2). We shall begin with the definition of uniform convergence for integrals with infinite limits of integration.

Definition 1. We say that integral (10.23) *is* ***uniformly convergent with respect to the parameter*** *y* ***on the interval*** *$c \leqslant y \leqslant d$ if, given an arbitrary $\varepsilon>0$, there is $L=L(\varepsilon)$ such that the inequality*

$$\left|J(y)-\int_a^l f(x, y)\,dx\right|=\left|\int_l^{+\infty} f(x, y)\,dx\right|<\varepsilon \tag{10.26}$$

is fulfilled for all $l>L(\varepsilon)$ and all $y \in [c, d]$ simultaneously.

Uniform convergence of an integral of an unbounded function is defined in a similar fashion:

Definition 2. Integral (10.25) *is said* ***to converge uniformly with respect to the parameter*** *y* ***on the interval*** *$[c, d]$ if for every $\varepsilon>0$ there exists $\delta=\delta(\varepsilon)>0$ such that the inequality*

$$\left|J^*(y)-\int_a^{b-\lambda} f(x, y)\,dx\right|=\left|\int_{b-\lambda}^{b} f(x, y)\,dx\right|<\varepsilon \tag{10.27}$$

holds for all $\lambda<b-a$ satisfying the condition $0<\lambda<\delta(\varepsilon)$ and for all $y \in [c, d]$ simultaneously.

Examples

1. The integral $J(y) = \int\limits_0^{+\infty} ye^{-xy}\,dx$ is convergent for every y belonging to the interval $0 \leqslant y \leqslant 1$ but not uniformly convergent. Indeed, we have

$$\int\limits_l^{+\infty} ye^{-xy}\,dx = \int\limits_{ly}^{+\infty} e^{-t}\,dt = e^{-ly}$$

Therefore, for an arbitrarily large fixed value $l > 0$ the latter integral exceeds $\frac{1}{2}$ for all values of y located sufficiently close to zero and, consequently, for $\varepsilon = \frac{1}{2}$ there is no $L(\varepsilon)$ such that for $l > L(\varepsilon)$ and for all y belonging to the interval $0 \leqslant y \leqslant 1$ the inequality

$$\left|\int\limits_l^{+\infty} ye^{-xy}\,dx\right| < \varepsilon = \frac{1}{2}$$

is fulfilled.

But if the interval $0 \leqslant y \leqslant 1$ is replaced by an interval of the form $0 < \delta \leqslant y \leqslant 1$, $\delta < 1$, the integral $J(y) = \int\limits_0^{+\infty} ye^{-xy}\,dx$ is uniformly convergent on the latter interval. In fact, we have $\int\limits_l^{+\infty} ye^{-xy}\,dx = \int\limits_{ly}^{+\infty} e^{-t}\,dt = e^{-ly} \leqslant e^{-l\delta}$ for $0 < \delta \leqslant y \leqslant 1$ and therefore the inequality

$$\left|\int\limits_l^{+\infty} ye^{-xy}\,dx\right| < \varepsilon$$

holds for $l > \frac{\ln\frac{1}{\varepsilon}}{\delta}$, $0 < \varepsilon < 1$, and all y belonging to the interval $0 < \delta \leqslant y \leqslant 1$.

2. The integral $J(y) = \int\limits_0^1 yx^{y-1}\,dx$ is convergent for every y belonging to the interval $0 \leqslant y \leqslant 1$ but not uniformly convergent. To show this, we take into account that here the integrand is unbounded for $x \to 0 + 0$. Let us estimate the integral $\int\limits_0^{\lambda} yx^{y-1}\,dx =$

$= x^y \Big|_0^\lambda = \lambda^y$. For any arbitrarily small and fixed $\lambda > 0$ this integral tends to unity as $y \to 0+0$. Hence, for $\varepsilon = \frac{1}{2}$, there is no $\delta = \delta(\varepsilon)$ such that the inequality $\left| \int_0^\lambda y x^{y-1}\, dx \right| < \varepsilon = \frac{1}{2}$ simultaneously holds for all y belonging to the interval $0 \leqslant y \leqslant 1$.

But if we replace the interval $0 \leqslant y \leqslant 1$ by an interval $0 < \delta_0 \leqslant \leqslant y \leqslant 1$, $\delta_0 < 1$, the integral $J(y) = \int_0^1 y x^{y-1}\, dx$ converges uniformly on the latter. Indeed, we have $\int_0^\lambda y x^{y-1}\, dx = \lambda^y \leqslant \lambda^{\delta_0}$ for $0 < \lambda < 1$ and $\delta_0 \leqslant y \leqslant 1$. Therefore, if $0 < \varepsilon < 1$ and $\lambda < \varepsilon^{\frac{1}{\delta_0}}$, we obtain the inequality $\left| \int_0^\lambda y x^{y-1}\, dx \right| < \varepsilon$ for all y belonging to the interval $\delta_0 \leqslant y \leqslant 1$.

2. Reducing Improper Integral Dependent on Parameter to a Functional Sequence. An improper integral dependent on a parameter can be reduced to a functional sequence which makes it possible to prove the fundamental theorems concerning such integrals on the basis of the corresponding theorems on functional sequences.

If an integral

$$J(y) = \int_a^{+\infty} f(x, y)\, dx \tag{10.28}$$

converges for every $y \in [c, d]$, then, for an arbitrary number sequence $l_1, l_2, \ldots, l_k, \ldots$, $\lim\limits_{k \to +\infty} l_k = +\infty$, where $l_k \geqslant a$ for $k = 1, 2, \ldots$, the functional sequence

$$F_k(y) = \int_a^{l_k} f(x, y)\, dx, \quad k = 1, 2, \ldots, \quad c \leqslant y \leqslant d$$

is obviously convergent to $J(y)$ on the interval $[c, d]$.

The theorem below holds under the condition that integral (10.28) converges for every y belonging to the interval $c \leqslant y \leqslant d$.

Theorem 10.5. *For the integral $J(y) = \int_a^{+\infty} f(x, y)\, dy$ to be uniformly convergent with respect to the parameter y on the interval*

$[c, d]$, *it is necessary and sufficient that the functional sequence*

$$F_k(y) = \int_a^{l_k} f(x, y)\,dx, \qquad k = 1, 2, \ldots \tag{10.29}$$

converge uniformly to $J(y)$ *on the interval* $c \leqslant y \leqslant d$ *for any choice of the sequence* $l_1, l_2, \ldots, l_k, \ldots, \lim\limits_{k\to+\infty} l_k = +\infty$.

Proof. Necessity. Suppose integral (10.28) is uniformly convergent on the interval $c \leqslant y \leqslant d$. Then, given an arbitrary $\varepsilon > 0$, there is $L(\varepsilon)$ such that, for all $l > L(\varepsilon)$, the inequality

$$\left| J(y) - \int_a^l f(x, y)\,dx \right| < \varepsilon$$

is satisfied for all $y \in [c, d]$ simultaneously.

Let $l_k \to +\infty$ for $k \to +\infty$ (where $l_k \geqslant a$, $k = 1, 2, \ldots$). Then, there exists $N(\varepsilon)$ such that $l_k > L(\varepsilon)$ for all $k \geqslant N(\varepsilon)$. Consequently, for all such k, we have, by the manner $L(\varepsilon)$ has been chosen, the inequality

$$|J(y) - F_k(y)| = \left| J(y) - \int_a^{l_k} f(x, y)\,dx \right| < \varepsilon$$

which holds for all $y \in [c, d]$. This means that sequence (10.29) is uniformly convergent to integral (10.28) on the interval $c \leqslant y \leqslant d$.

Sufficiency. Let us show that if every sequence of form (10.29) where $\lim\limits_{k\to+\infty} l_k = +\infty$ converges uniformly to $J(y)$ on the interval $c \leqslant y \leqslant d$, integral (10.28) is uniformly convergent with respect to the parameter y on this interval. In fact, if we suppose that integral (10.28) (which is, by the hypothesis, convergent for every $y \in [c, d]$) converges nonuniformly with respect to y on the interval $c \leqslant y \leqslant d$, there must exist ε_0 such that for an arbitrarily large L there are $l > L$ and $y \in [c, d]$ such that the inequality

$$\left| J(y) - \int_a^l f(x, y)\,dx \right| \geqslant \varepsilon_0$$

is satisfied. Then, making L assume the values $L = 1, 2, 3, \ldots, k, \ldots$ we obtain the corresponding number sequence l_k, $k = 1, 2, \ldots$, $l_k > k$, and a sequence $y_k \in [c, d]$, $k = 1, 2, \ldots$, for which

$$\left| J(y_k) - \int_a^{l_k} f(x, y_k)\,dx \right| = |J(y_k) - F_k(y_k)| \geqslant \varepsilon_0.$$

This means that the functional sequence $F_k(y) = \int\limits_a^{l_k} f(x, y)\,dx$, $k = 1, 2, \ldots$, thus constructed, converges nonuniformly on the interval $c \leqslant y \leqslant d$, which contradicts the hypothesis. The theorem has been proved.

Note 1. If $f(x, y)$ is a function, retaining its sign, for instance, a nonnegative one, then for the integral $J(y) = \int\limits_a^{+\infty} f(x, y)\,dx$ to be uniformly convergent with respect to y on the interval $c \leqslant y \leqslant d$, it is sufficient that functional sequence (10.29) converge to $J(y)$ uniformly for at least one particular choice of the number sequence $l_1, l_2, \ldots, l_k, \ldots$.

Indeed, if $f(x, y)$ is nonnegative we have $\int\limits_a^{l} f(x, y)\,dx \geqslant \int\limits_a^{l_k} f(x, y)\,dx$ for all $l \geqslant l_k$. Consequently,

$$\left|J(y) - \int\limits_a^{l} f(x, y)\,dx\right| \leqslant \left|J(y) - \int\limits_a^{l_k} f(x, y)\,dx\right| < \varepsilon \qquad \text{for all } l > l_k$$

and for all $y \in [c, d]$ simultaneously provided that l_k is sufficiently large.

Note 2. If the function $f(x, y)$ is continuous for $a \leqslant x < +\infty$, $c \leqslant y \leqslant d$ and retains its sign (for instance, is nonnegative) and the integral $\int\limits_a^{+\infty} f(x, y)\,dx$ is a continuous function of the parameter y on the interval $[c, d]$, this integral is uniformly convergent on $[c, d]$.

In fact, taking an increasing number sequence $l_1, l_2, \ldots, l_k, \ldots$ $\ldots$, $\lim\limits_{k\to+\infty} l_k = +\infty$ $(l_k \geqslant a,\ k = 1, 2, \ldots)$ we arrive at the functional sequence

$$F_k(y) = \int\limits_a^{l_k} f(x, y)\,dx, \qquad k = 1, 2, \ldots \tag{A}$$

The function $f(x, y)$ being nonnegative, sequence (A) is monotone nondecreasing, and, by Theorem 10.1, the functions $F_k(y)$, $k = 1, 2, \ldots$, are continuous. Besides, this sequence converges to the

continuous function

$$J(y)=\int_a^{+\infty} f(x, y)\,dx \tag{B}$$

on the interval $c \leqslant y \leqslant d$. But then Dini's theorem (see Chapter 8, § 2, Sec. 1) implies that sequence (A) converges uniformly to its limit (B) on the interval $[c, d]$ and, consequently, by Note 1, the integral $J(y)=\int_a^{+\infty} f(x, y)\,dx$ is uniformly convergent on this interval.

Note 3. The improper integral

$$J^*(y)=\int_a^b f(x, y)\,dx=\lim_{\lambda\to 0+0}\int_a^{b-\lambda} f(x, y)\,dx$$

can be similarly reduced to a functional sequence $F_k^*(y)=$ $=\int_a^{b-\lambda_k} f(x, y)\,dx$ where $\lambda_k \to 0+0$ for $k\to+\infty$ but we shall not go into particulars here.

3. Properties of Uniformly Convergent Improper Integrals Dependent on Parameter.

Theorem 10.6. *If $f(x, y)$ is a continuous function defined in a domain $a \leqslant x < +\infty$, $c \leqslant y \leqslant d$, and the integral $J(y)=$ $=\int_a^{+\infty} f(x, y)\,dx$ converges uniformly with respect to the parameter y on the interval $c \leqslant y \leqslant d$, the function $J(y)$ is continuous on this interval.*

Proof. Take an arbitrary numerical sequence $l_1, l_2, \ldots, l_k, \ldots$ $\ldots$, $\lim\limits_{k\to+\infty} l_k \to +\infty$, $(l_k \geqslant a)$, and consider the sequence of functions

$$F_k(y)=\int_a^{l_k} f(x, y)\,dx,\quad k=1, 2, \ldots,\quad c\leqslant y\leqslant d$$

By Theorem 10.1 on the continuous dependence of a proper integral on the parameter, the functions $F_k(y)$, $k=1, 2, \ldots$ are continuous on the interval $c\leqslant y\leqslant d$. But Theorem 10.5 implies that this sequence is uniformly convergent to the integral $J(y)=$ $=\int_a^{+\infty} f(x, y)\,dx$ on the interval $c\leqslant y\leqslant d$, and consequently the

function $J(y)$ is continuous as the limit of a uniformly convergent functional sequence. The theorem has been proved.

Theorem 10.7 (On Differentiation of an Improper Integral with Respect to the Parameter). *Let $f(x, y)$ and $f'_y(x, y)$ be continuous for $c \leqslant y \leqslant d$, $a \leqslant x < +\infty$, and let the integral*

$$J(y) = \int_a^{+\infty} f(x, y)\,dx \tag{10.28'}$$

be convergent on the interval $c \leqslant y \leqslant d$. Suppose that the integral

$$\int_a^{+\infty} f'_y(x, y)\,dx \tag{10.30}$$

converges uniformly on this interval. Then $J(y)$ is a differentiable function of y for $y \in [c, d]$ and

$$\frac{dJ}{dy} = \frac{d}{dy}\int_a^{+\infty} f(x, y)\,dx = \int_a^{+\infty} f'_y(x, y)\,dx \tag{10.31}$$

Proof. Take an arbitrary number sequence $l_1, l_2, \ldots, l_k, \ldots$, $\lim\limits_{k\to+\infty} l_k = +\infty$, $l_k \geqslant a$, and consider the sequence of functions

$$F_k(y) = \int_a^{l_k} f(x, y)\,dx, \quad k = 1, 2, \ldots, c \leqslant y \leqslant d$$

which is convergent to the integral $J(y) = \int\limits_a^{+\infty} f(x, y)\,dx$ on the interval $[c, d]$. Theorem 10.3 on differentiating a proper integral with respect to the parameter implies

$$F'_k(y) = \frac{d}{dy}\int_a^{l_k} f_y(x, y)\,dx = \int_a^{l_k} f'_y(x, y)\,dx, \quad k = 1, 2, \ldots,$$
$$c \leqslant y \leqslant d$$

where all the functions $F'_k(y)$, $k = 1, 2, \ldots$, are continuous on $[c, d]$. Integral (10.30) converging uniformly, the sequence of functions $F'_k(y)$ is uniformly convergent to integral (10.30) on the interval $[c, d]$. Hence, we have

$$F_k(y) \to J(y) \quad \text{on } [c, d], \qquad F'_k(y) \rightrightarrows \int_a^{+\infty} f'_y(x, y)\,dx \quad \text{on } [c, d]$$

where the symbol $\rightrightarrows$ indicates uniform convergence, and $F_k'(y)$ are continuous on $[c, d]$. Consequently, by the theorem on differentiating a functional sequence (see Sec. 3 in § 2 of Chapter 8), $J(y)$ is a differentiable function on $[c, d]$ and the relation

$$J'(y) = \int_a^{+\infty} f_y'(x, y)\,dx \tag{10.31}$$

holds for every y belonging to the interval $c \leqslant y \leqslant d$ which is what we set out to prove.

Theorem 10.8 (On Integration of an Improper Integral with Respect to the Parameter). *If $f(x, y)$ is a continuous function in the domain $a \leqslant x < +\infty$, $c \leqslant y \leqslant d$, and the integral*

$$J(y) = \int_a^{+\infty} f(x, y)\,dx \tag{10.28''}$$

converges uniformly on the interval $c \leqslant y \leqslant d$, then

$$\int_c^d J(y)\,dy = \int_c^d dy \int_a^{+\infty} f(x, y)\,dx = \int_a^{+\infty} dx \int_c^d f(x, y)\,dy \tag{10.32}$$

Proof. For any number sequence $l_1, l_2, \ldots, l_k, \ldots$ ($l_k \geqslant a$, $\lim\limits_{k\to\infty} l_k = +\infty$), the corresponding sequence of functions

$$F_k(y) = \int_a^{l_k} f(x, y)\,dx, \qquad k = 1, 2, \ldots$$

is uniformly convergent to $J(y)$ on $[c, d]$ which is implied by Theorem 10.5 on reducing a uniformly convergent integral to the corresponding functional sequence. By Theorem 10.1 on the continuity of a proper integral as a function of its parameter, all $F_k(y)$, $k = 1, 2, \ldots$, are continuous functions on the interval $c \leqslant y \leqslant d$. Hence, by the theorem on integrating a functional sequence (see Chapter 8, § 2, Sec. 2), we have $\lim\limits_{k\to+\infty} \int_c^d F_k(y)\,dy = \int_c^d J(y)\,dy$. But the theorem on integration of a proper integral with respect to the parameter implies that

$$\int_c^d F_k(y)\,dy = \int_c^d dy \int_a^{l_k} f(x, y)\,dx = \int_a^{l_k} dx \int_c^d f(x, y)\,dy$$

Consequently, for any choice of the number sequence $l_1, l_2, \ldots$ $\ldots, l_k, \ldots, \lim\limits_{k\to+\infty} l_k = +\infty$, we have

$$\lim_{k\to+\infty} \int_a^{l_k} dx \int_c^d f(x, y)\,dy = \int_c^d J(y)\,dy$$

This means that the integral $\int\limits_a^{+\infty} dx \int\limits_c^d f(x, y)\,dy$ converges and the equality

$$\int_a^{+\infty} dx \int_c^d f(x, y)\,dy = \int_c^d dy \int_a^{+\infty} f(x, y)\,dx$$

is fulfilled. The theorem has been proved.

Corollary. *If $f(x, y)$ is a continuous function retaining its sign and defined for $a \leqslant x < +\infty$, $c \leqslant y \leqslant d$ (e.g. $f(x, y)$ is nonnegative) and if the integral*

$$J(y) = \int_a^{+\infty} f(x, y)\,dx$$

is a continuous function of y for $c \leqslant y \leqslant d$, relation (10.32) *is valid.*

Proof. In fact, Note 2 after Theorem 10.5 implies that the integral $J(y) = \int\limits_a^{+\infty} f(x, y)\,dx$ converges uniformly on the interval $c \leqslant y \leqslant d$ and, consequently, by Theorem 10.8, equality (10.32) is valid.

For a function $f(x, y)$ retaining its sign we can prove the following

Theorem 10.9 (On Reversing the Order of Integration in an Improper Iterated Integral). *Let $f(x, y)$ be a continuous function of constant sign defined for $c \leqslant y < +\infty$, $a \leqslant x < +\infty$, and let the integrals*

$$J(y) = \int_a^{+\infty} f(x, y)\,dx \quad \textit{and} \quad J^*(x) = \int_c^{+\infty} f(x, y)\,dy$$

regarded as functions of the corresponding parameters be, respectively, continuous for $c \leqslant y < +\infty$ and $a \leqslant x < +\infty$. Then, if at least one of the iterated integrals

$$\int_c^{+\infty} dy \int_a^{+\infty} f(x, y)\,dx \quad \textit{and} \quad \int_a^{+\infty} dx \int_c^{+\infty} f(x, y)\,dy$$

converges, the other integral also converges and their values coincide.

Proof. We shall present the proof for the case of a nonnegative function $f(x, y)$ defined for $c \leqslant y < +\infty$, $a \leqslant x < +\infty$. Let us suppose that the iterated integral

$$J = \int_c^{+\infty} dy \int_a^{+\infty} f(x, y)\, dx \tag{10.33}$$

is convergent. Then we must prove that

$$\lim_{l \to +\infty} \int_a^l dx \int_c^{+\infty} f(x, y)\, dy = J = \int_c^{+\infty} dy \int_a^{+\infty} f(x, y)\, dx \tag{10.34}$$

Let there be given $\varepsilon > 0$. We shall show that, for all sufficiently large l, the difference between the variable quantity $\int_a^l dx \int_c^{+\infty} f(x, y)\, dy$ and the constant quantity $\int_a^{+\infty} dy \int_c^{+\infty} f(x, y)\, dx$ is less than ε in its absolute value.

We have, by the Corollary of Theorem 10.8, the relation

$$\int_a^l dx \int_c^{+\infty} f(x, y)\, dy = \int_c^{+\infty} dy \int_a^l f(x, y)\, dx$$

The function $f(x, y)$ being nonnegative, we can write

$$\begin{aligned}
0 &\leqslant \int_c^{+\infty} dy \int_a^{+\infty} f(x, y)\, dx - \int_a^l dx \int_c^{+\infty} f(x, y)\, dy = \\
&= \int_c^{+\infty} dy \int_l^{+\infty} f(x, y)\, dx = \\
&= \int_c^{c_1} dy \int_l^{+\infty} f(x, y)\, dx + \int_{c_1}^{+\infty} dy \int_l^{+\infty} f(x, y)\, dx \leqslant \\
&\leqslant \int_c^{c_1} dy \int_l^{+\infty} f(x, y)\, dx + \int_{c_1}^{+\infty} dy \int_a^{+\infty} f(x, y)\, dx
\end{aligned} \tag{10.35}$$

where $c < c_1 < +\infty$. We begin with estimating the second summand on the right-hand side of inequality (10.35). Since iterated integral (10.33) is convergent, there is $c_1 > c$ such that

$$\int_{c_1}^{+\infty} dy \int_a^{+\infty} f(x, y)\, dx < \frac{\varepsilon}{2} \tag{10.36}$$

Let us fix $c_1 > c$ in such a way that inequality (10.36) holds and then proceed to estimate the first summand on the right-hand side of inequality (10.35). By the hypothesis, the integral $\int\limits_a^{+\infty} f(x, y)\,dx$ is a continuous function for $c \leqslant y < +\infty$ and, consequently, it converges uniformly because the function $f(x, y)$ is nonnegative (see Note 2 after Theorem 10.5). Therefore there exists $L(\varepsilon)$ such that the inequality $\int\limits_l^{+\infty} f(x, y)\,dx < \frac{\varepsilon}{2(c_1-c)}$ is fulfilled for all $l > L(\varepsilon)$ and for all $y \in [c, c_1]$ simultaneously. But then we have

$$\int\limits_c^{c_1} dy \int\limits_l^{+\infty} f(x, y)\,dx < \frac{\varepsilon(c_1-c)}{2(c_1-c)} = \frac{\varepsilon}{2} \tag{10.37}$$

for all $l > L(\varepsilon)$. Now, taking into account (10.35), (10.36) and (10.37) we conclude that

$$0 \leqslant \int\limits_c^{+\infty} dy \int\limits_a^{+\infty} f(x, y)\,dx - \int\limits_a^{l} dx \int\limits_c^{+\infty} f(x, y)\,dy < \varepsilon$$

for all $l > L(\varepsilon)$ which is what we set out to prove.

If $f(x, y)$ is a function that may have alternating sign the theorem on reversing the order of integration in an improper iterated integral can be formulated as follows.

Theorem 10.9'. *Let the function $f(x, y)$ be continuous for $a \leqslant x < +\infty$, $c \leqslant y < +\infty$ and let the integrals*

$$\int\limits_c^{+\infty} f(x, y)\,dy \quad \text{and} \quad \int\limits_a^{+\infty} f(x, y)\,dx \tag{10.38}$$

be, respectively, uniformly convergent on every finite interval $a \leqslant x \leqslant A$ and on every finite interval $c \leqslant y \leqslant C$. Then, if at least one of the iterated integrals

$$\int\limits_a^{+\infty} dx \int\limits_c^{+\infty} |f(x, y)|\,dy \quad \text{and} \quad \int\limits_c^{+\infty} dy \int\limits_a^{+\infty} |f(x, y)|\,dx \tag{10.39}$$

converges, the iterated integrals

$$\int\limits_a^{+\infty} dx \int\limits_c^{+\infty} f(x, y)\,dy \qquad \text{and} \quad \int\limits_c^{+\infty} dy \int\limits_a^{+\infty} f(x, y)\,dx \tag{10.40}$$

are also convergent and their values are equal.

Proof. For definiteness, suppose that the second integral (10.39) converges. Then, applying the comparison test to the functions $f(x, y)$ and $|f(x, y)|$ and to the functions $\int\limits_a^{+\infty} f(x, y)\,dx$ and $\int\limits_a^{+\infty} |f(x, y)|\,dx$ we conclude that the second integral (10.40) also converges.

Thus, we must only prove that

$$\lim_{l\to+\infty}\int\limits_a^l dx\int\limits_c^{+\infty} f(x, y)\,dy=\int\limits_c^{+\infty} dy\int\limits_a^{+\infty} f(x, y)\,dx \tag{10.41}$$

The integral $\int\limits_c^{+\infty} f(x, y)\,dy$ being uniformly convergent, we have

$$\int\limits_a^l dx\int\limits_c^{+\infty} f(x, y)\,dy=\int\limits_c^{+\infty} dy\int\limits_a^l f(x, y)\,dx \tag{10.42}$$

for every finite $l > a$. Let us estimate the difference between the variable quantity $\int\limits_a^l dx\int\limits_c^{+\infty} f(x, y)\,dy$ and the constant quantity $\int\limits_c^{+\infty} dy\int\limits_a^{+\infty} f(x, y)\,dx$ entering into relation (10.41). Taking advantage of equality (10.42), we see that

$$\begin{aligned}\left|\int\limits_c^{+\infty} dy\int\limits_a^{+\infty} f(x, y)\,dx-\int\limits_a^l dx\int\limits_c^{+\infty} f(x, y)\,dy\right|&=\\
=\left|\int\limits_c^{+\infty} dy\int\limits_a^{+\infty} f(x, y)\,dx-\int\limits_c^{+\infty} dy\int\limits_a^l f(x, y)\,dx\right|&=\\
=\left|\int\limits_c^{+\infty} dy\int\limits_l^{+\infty} f(x, y)\,dx\right|=\left|\int\limits_c^{c_1} dy\int\limits_l^{+\infty} f(x, y)\,dx\right.&+\\
\left.+\int\limits_{c_1}^{+\infty} dy\int\limits_l^{+\infty} f(x, y)\,dx\right|\leqslant\left|\int\limits_c^{c_1} dy\int\limits_l^{+\infty} f(x, y)\,dx\right|&+\\
+\int\limits_{c_1}^{+\infty} dy\int\limits_l^{+\infty} |f(x, y)|\,dx\leqslant\left|\int\limits_c^{c_1} dy\int\limits_l^{+\infty} f(x, y)\,dx\right|+\int\limits_{c_1}^{+\infty} dy\int\limits_a^{+\infty} |f(x, y)|\,dx&\end{aligned} \tag{10.43}$$

for any $c_1 > c$. Since, by the hypothesis, the iterated integral $\int\limits_c^{+\infty} dy \int\limits_a^{+\infty} |f(x, y)|\, dx$ converges, for every $\varepsilon > 0$ there is $c_1 > c$ such that

$$\int\limits_{c_1}^{+\infty} dy \int\limits_a^{+\infty} |f(x, y)|\, dx < \frac{\varepsilon}{2} \tag{10.44}$$

Now, fixing a value of $c_1 > c$ for which inequality (10.44) holds and taking into account that the integral $\int\limits_a^{+\infty} f(x, y)\, dx$ is uniformly convergent, we choose, as in the proof of Theorem 10.9, a quantity $L(\varepsilon)$ such that inequality $\left|\int\limits_l^{+\infty} f(x, y)\, dx\right| < \frac{\varepsilon}{2(c_1 - c)}$ is fulfilled for all $l > L(\varepsilon)$ and for all $y \in [c, c_1]$. Then we have

$$\left|\int\limits_c^{c_1} dy \int\limits_l^{+\infty} f(x, y)\, dx\right| < \frac{\varepsilon (c_1 - c)}{2(c_1 - c)} = \frac{\varepsilon}{2} \tag{10.45}$$

for all $l > L(\varepsilon)$ and, consequently, by virtue of (10.43), (10.44) and (10.45), we obtain the inequality

$$\left|\int\limits_c^{+\infty} dy \int\limits_a^{+\infty} f(x, y)\, dx - \int\limits_a^{l} dx \int\limits_c^{+\infty} f(x, y)\, dy\right| < \varepsilon$$

for all such l, which is what we set out to prove.

It should be noted that similar theorems are also valid for improper integrals of unbounded functions involving a parameter.

4. Tests for Uniform Convergence of Improper Integrals Dependent on Parameter.

Cauchy's Test. For an integral $\int\limits_a^{+\infty} f(x, y)\, dx$ *to converge uniformly on an interval* $[c, d]$ *it is necessary and sufficient that for every* $\varepsilon > 0$ *there exist* $L = L(\varepsilon)$ *such that the inequality*

$$\left|\int\limits_{l'}^{l''} f(x, y)\, dx\right| < \varepsilon \tag{10.46}$$

holds for all $l', l'' > L(\varepsilon)$ *and for all* $y \in [c, d]$ *simultaneously.*

Proof. Necessity. If the integral is uniformly convergent, then, for every $\varepsilon > 0$ there is $L = L(\varepsilon)$ such that for all $l'_1 > L(\varepsilon)$,

$l'' > L(\varepsilon)$ and $y \in [c, d]$ the inequalities

$$\left|\int_{l'}^{+\infty} f(x, y)\,dx\right| < \frac{\varepsilon}{2} \quad \text{and} \quad \left|\int_{l''}^{+\infty} f(x, y)\,dx\right| < \frac{\varepsilon}{2}$$

are fulfilled. Therefore, for all l', $l'' > L(\varepsilon)$ and all $y \in [c, d]$ we obtain the inequality

$$\left|\int_{l'}^{l''} f(x, y)\,dx\right| = \left|\int_{l'}^{+\infty} f(x, y)\,dx - \int_{l''}^{+\infty} f(x, y)\,dx\right| \leqslant$$

$$\leqslant \left|\int_{l'}^{+\infty} f(x, y)\,dx\right| + \left|\int_{l''}^{+\infty} f(x, y)\,dx\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Sufficiency. If inequality (10.46) holds for all $l' > L(\varepsilon)$, $l'' > L(\varepsilon)$ and all $y \in [c, d]$, the integral $\int_{a}^{+\infty} f(x, y)\,dx$ converges for every $y \in [c, d]$ (see Chapter 9, § 1, Sec. 3). Therefore, passing to the limit as $l'' \to +\infty$ we obtain, for all $l' > L(\varepsilon)$, the inequality

$$\left|\int_{l'}^{+\infty} f(x, y)\,dx\right| \leqslant \varepsilon$$

which holds for all $y \in [c, d]$ simultaneously. The theorem has been proved.

Weierstrass' Test (Sufficient Condition for Uniform Convergence of an Improper Integral). *If $|f(x, y)| \leqslant g(x)$ for $a \leqslant x < +\infty$ and the integral $\int_{a}^{+\infty} g(x)\,dx$ is convergent, the integrals $\int_{a}^{+\infty} f(x, y)\,dx$ and $\int_{a}^{+\infty} |f(x, y)|\,dx$ are uniformly convergent on the interval $c \leqslant y \leqslant d$.*

Proof. Let an arbitrary $\varepsilon > 0$ be given. The integral $\int_{a}^{+\infty} g(x)\,dx$ converging, there is $L = L(\varepsilon)$ such that the condition

$$\int_{l'}^{l''} g(x)\,dx < \varepsilon$$

is satisfied for all l', $l'' > L(\varepsilon)$ $(l'' > l')$. This implies that, for all l', $l'' > L(\varepsilon)$ $(l'' > l')$, the inequalities

$$\left|\int_{l'}^{l''} f(x, y)\,dx\right| \leqslant \int_{l'}^{l''} |f(x, y)|\,dx \leqslant \int_{l'}^{l''} g(x)\,dx < \varepsilon \qquad (l'' > l')$$

are fulfilled for all $y \in [c, d]$ simultaneously. Consequently, by Cauchy's test, the integrals $\int\limits_a^{+\infty} f(x, y)\,dx$ and $\int\limits_a^{+\infty} |f(x, y)|\,dx$ are uniformly convergent on the interval $c \leqslant y \leqslant d$ which is what we set out to prove.

The corresponding tests for uniform convergence of improper integrals with unbounded integrands and finite limits of integration are formulated and proved in a similar way. As an instance, let us formulate

Cauchy's Test (Necessary and Sufficient Condition for Uniform Convergence of an Improper Integral of an Unbounded Function). *An improper integral of the form*

$$J^*(y) = \int_a^b f(x, y)\,dx = \lim_{\lambda \to 0+0} \int_a^{b-\lambda} f(x, y)\,dx, \qquad c \leqslant y \leqslant d \qquad (10.47)$$

is uniformly convergent with respect to the parameter y on the interval $c \leqslant y \leqslant d$ if and only if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all λ' and λ'' belonging to the interval $0 < \lambda < \min(b - a, \delta(\varepsilon))$ the inequality

$$\left|\int_{b-\lambda'}^{b-\lambda''} f(x, y)\,dx\right| < \varepsilon \qquad (10.48)$$

is fulfilled for all $y \in [c, d]$ simultaneously.

Examples

1. It is clear that the integral $J(p) = \int\limits_0^1 x^{p-1}\,dx$ is convergent for $p > 0$ and divergent for $p \leqslant 0$. Let $p_0 > 0$. Then the inequality $x^{p-1} \leqslant x^{p_0-1}$ holds on the interval $0 < x < 1$ for all $p \geqslant p_0$. Therefore, we can put $f(x, p) = x^{p-1}$ and $g(x) = x^{p_0-1}$ and apply Weierstrass' test and thus conclude that, since the integral

$$\int_0^1 g(x)\,dx = \int_0^1 x^{p_0-1}\,dx = \frac{x^{p_0}}{p_0}\Bigg|_0^1 = \frac{1}{p_0}$$

converges, the integral $J(p)=\int_0^1 f(x,p)\,dx=\int_0^1 x^{p-1}\,dx$ converges uniformly with respect to the parameter p on the interval $0<p_0\leqslant p<+\infty$ for every arbitrarily small and fixed $p_0>0$.

Let us test this integral for uniform convergence on the interval $0<p<+\infty$. For this purpose we shall study the behaviour of the integral $\int_0^\lambda x^{p-1}\,dx$ for $p\to 0+0$. We have $\int_0^\lambda x^{p-1}\,dx=\frac{\lambda^p}{p}\to +\infty$ for $p\to 0+0$ and for any arbitrarily small fixed $\lambda>0$. Consequently, for any $\varepsilon>0$ and any arbitrarily small $\lambda>0$, the inequality

$$\left|\int_0^\lambda x^{p-1}\,dx\right|<\varepsilon$$

cannot be valid for all p belonging to the interval $0<p<+\infty$. This means that the integral $J(p)=\int_0^1 x^{p-1}\,dx$ converges nonuniformly on the interval $0<p<+\infty$.

2. The integral $J(\alpha)=\int_0^{+\infty} e^{-\alpha x^2}\,dx$ is uniformly convergent for $0<\alpha_0\leqslant\alpha<+\infty$ where $\alpha_0>0$ is an arbitrarily small fixed positive number. This is implied by Weierstrass' test if we put $f(x,\alpha)=e^{-\alpha x^2}$ and $g(x)=e^{-\alpha_0 x^2}$ since we have $|f(x,\alpha)|=e^{-\alpha x^2}\leqslant g(x)=e^{-\alpha_0 x^2}$ for $0<\alpha_0\leqslant\alpha<+\infty$, $0\leqslant x<+\infty$, and the integral $\int_0^{+\infty} g(x)\,dx=\int_0^{+\infty} e^{-\alpha_0 x^2}\,dx$ is converging.

Let us prove that if we substitute the interval $0<\alpha<+\infty$ for the interval $0<\alpha_0\leqslant\alpha<+\infty$ and thus consider the values of α which can be arbitrarily close to zero, i.e. take the maximal range of α for which the integral $J(\alpha)=\int_0^{+\infty} e^{-\alpha x^2}\,dx$ is convergent, the uniform convergence of the integral is violated.

To show that the integral under consideration is not uniformly convergent on the interval $0<\alpha<+\infty$, we take into account that $\int_0^{+\infty} e^{-t^2}\,dt$ is a positive constant as an integral of a nonnegative continuous function which is not identically equal to zero. This

enables us to estimate the integral $\int\limits_l^{+\infty} e^{-\alpha x^2}\,dx$ for an arbitrarily large fixed l and $0<\alpha<+\infty$. Putting $t=x\sqrt{\alpha}$ ($dt=\sqrt{\alpha}\,dx$) we obtain

$$\int\limits_l^{+\infty} e^{-\alpha x^2}\,dx=\frac{1}{\sqrt{\alpha}}\int\limits_{l\sqrt{\alpha}}^{+\infty} e^{-t^2}\,dt\to+\infty \quad \text{for } \alpha\to 0+0$$

since

$$\lim_{\alpha\to 0+0}\int\limits_{l\sqrt{\alpha}}^{+\infty} e^{-t^2}\,dt=\int\limits_0^{+\infty} e^{-t^2}\,dt=\text{const}>0$$

Hence, for any $l>0$ and any arbitrarily small fixed $\varepsilon>0$ the inequality

$$\left|\int\limits_l^{+\infty} e^{-\alpha x^2}\,dx\right|=\left|\frac{1}{\sqrt{\alpha}}\int\limits_{l\sqrt{\alpha}}^{+\infty} e^{-t^2}\,dt\right|<\varepsilon$$

cannot be valid simultaneously for all α belonging to the interval $0<\alpha<+\infty$, which means that the integral does not converge uniformly on the whole interval $0<\alpha<+\infty$.

3. Let us prove that the integral $J(\alpha)=\int\limits_0^{+\infty} e^{-\alpha x}\frac{\sin\beta x}{x}\,dx$ is uniformly convergent with respect to the parameter α for $0\leqslant\alpha<<+\infty$ and any fixed $\beta\neq 0$. To do this, let us estimate the integral $\int\limits_l^{+\infty} e^{-\alpha x}\frac{\sin\beta x}{x}\,dx$, $l>0$. Putting $u=\frac{1}{x}$, $dv=e^{-\alpha x}\sin\beta x\,dx$ and integrating by parts we obtain

$$\int\limits_l^{+\infty} e^{-\alpha x}\frac{\sin\beta x}{x}\,dx=-\frac{e^{-\alpha x}\sin(\beta x+\varphi)}{x\sqrt{\alpha^2+\beta^2}}\Bigg|_{x=l}^{+\infty}-\int\limits_l^{+\infty}\frac{e^{-\alpha x}\sin(\beta x+\varphi)\,dx}{x^2\sqrt{\alpha^2+\beta^2}}$$

where φ is an auxiliary angle determined by the relations $\cos\varphi=\frac{\alpha}{\sqrt{\alpha^2+\beta^2}}$ and $\sin\varphi=\frac{\beta}{\sqrt{\alpha^2+\beta^2}}$. We have $\left|\frac{e^{-\alpha x}\sin\beta x}{\sqrt{\alpha^2+\beta^2}}\right|\leqslant\leqslant\frac{1}{|\beta|}$ for $\beta\neq 0$, all $x\geqslant 0$ and all $\alpha\geqslant 0$ and, consequently, the inequality

$$\left|\int\limits_l^{+\infty} e^{-\alpha x}\frac{\sin\beta x}{x}\,dx\right|\leqslant\frac{1}{l\,|\beta|}+\frac{1}{|\beta|}\int\limits_l^{+\infty}\frac{dx}{x^2}=$$

$$=\frac{2}{l\,|\beta|}<\varepsilon \quad \text{for } l>\frac{2}{|\beta|\,\varepsilon}$$

is fulfilled for all α, $0 \leqslant \alpha < +\infty$, which means that the integ
in question is uniformly convergent.

If we have an improper integral dependent on a parameter $y \in [c, d]$, whose one or both limits of integration are infinite a
whose integrand has one or more singular points, we break up t
interval of integration (provided this is possible) in such a w
that the integral taken over each part has either one infinite lin
of integration or one singular point. Then the original integral
said to be *uniformly convergent with respect to the parameter* y
the interval $c \leqslant y \leqslant d$ if and only if each of the constituent integr
taken over the parts the original interval of integration is divid
into is uniformly convergent for $c \leqslant y \leqslant d$.

5. Examples of Evaluating Improper Integrals Dependent on Par meter by Means of Differentiation and Integration with Respe to Parameter. The integrals below not only demonstrate some tec niques, but are also used in various divisions of mathematics a physics.

1. Knowing that

$$\int_0^{+\infty} \frac{x^{2m}}{1+x^{2n}}\,dx = \frac{\pi}{2n}\,\frac{1}{\sin\dfrac{2m+1}{2n}\pi} \qquad (\,$$

for $m < n$ where m and n are positive integers (see Example 3 i § 3 of Chapter 9), we shall prove, on the basis of the theorem c continuous dependence of an integral on a parameter, that

$$\int_0^{+\infty} \frac{t^{p-1}}{1+t}\,dt = \frac{\pi}{\sin p\pi} \qquad (\,$$

for $0 < p < 1$. For this purpose we substitute $x = t^{\frac{1}{2n}}$ into (and thus obtain

$$\int_0^{+\infty} \frac{t^{\frac{2m+1}{2n}-1}}{1+t}\,dt = \frac{\pi}{\sin\dfrac{2m+1}{2n}\pi} \qquad (1$$

The function $f(t, p) = \dfrac{t^{p-1}}{1+t}$ is continuous for $0 < t < +\infty$
$0 < p < 1$. Breaking up the interval of integration $0 \leqslant t < +\infty$ into two parts (e.g. $0 \leqslant t \leqslant 1$ and $1 \leqslant t < +\infty$) and applyin Weierstrass' test to the integrals over $[0, 1]$ and $[1, +\infty)$ with th function $g(t)$ equal, respectively, to $\dfrac{t^{p_1-1}}{1+t}$ and $\dfrac{t^{p_2-1}}{1+t}$, $0 < p_1 \leqslant$
$\leqslant p_2 < 1$, we see that the integral $\int\limits_0^{+\infty} \dfrac{t^{p-1}}{1+t}\,dt$ is uniformly con

vergent with respect to p on every interval of the form $0 < p_1 \leqslant \leqslant p \leqslant p_2 < 1$. Hence, the integral $\int\limits_0^{+\infty} \frac{t^{p-1}}{1+t}\,dt$ is a continuous function of the parameter p for $0 < p < 1$. Every number p belonging to the interval (0, 1) can be regarded as the limit of a subsequence of the number sequence $\frac{2m+1}{2n}$, $0 < m < n$, $m, n = 1, 2, \ldots$, and therefore, performing an appropriate passage to the limit in relation (B), we arrive at relation (A), which we set out to prove. Equality (A) is used in the theory of Euler's integrals (see § 4 of the present chapter).

2. Let us evaluate the integral $\int\limits_0^{+\infty} \frac{\sin \beta x}{x}\,dx$. It cannot be differentiated directly with respect to the parameter β under the integral sign but we know (see Sec. 2 of § 1) that the more general integral $\int\limits_0^{+\infty} e^{-\alpha x}\frac{\sin \beta x}{x}\,dx$ differing from the above integral in the factor $e^{-\alpha x}$, $\alpha > 0$ (which guarantees uniform convergence of the latter integral) can be evaluated by means of differentiation with respect to the parameter. This results in

$$\int\limits_0^{+\infty} e^{-\alpha x}\frac{\sin \beta x}{x}\,dx = \arctan\frac{\beta}{\alpha}$$

As was proved (see Example 3 in Sec. 4), this integral is uniformly convergent with respect to α, $0 \leqslant \alpha < +\infty$, for any fixed β. Consequently, it is a continuous function of the parameter α for $0 \leqslant \leqslant \alpha < +\infty$. Therefore,

$$\int\limits_0^{+\infty} \frac{\sin \beta x}{x}\,dx = \lim_{\alpha\to 0+0} \int\limits_0^{+\infty} e^{-\alpha x}\frac{\sin \beta x}{x}\,dx =$$

$$= \lim_{\alpha\to 0+0} \arctan\frac{\beta}{\alpha} = \begin{cases} \frac{\pi}{2} & \text{for } \beta > 0 \\ 0 & \text{for } \beta = 0 \\ -\frac{\pi}{2} & \text{for } \beta < 0 \end{cases} \tag{10.49}$$

In particular, we obtain

$$\int\limits_0^{+\infty} \frac{\sin x}{x}\,dx = \frac{\pi}{2} \tag{10.50}$$

The last integral is used in the theory of Fourier's series and integrals.

3. Let us evaluate the Euler-Poisson integral

$$J=\int_{0}^{+\infty} e^{-x^2}\,dx \tag{10.51}$$

(see also the end of Chapter 9). Its convergence was established in Chapter 9, § 1, Sec. 4. Putting $x = ut$, $dx = u\,dt$, we obtain

$$J=\int_{0}^{+\infty} e^{-u^2t^2}u\,dt$$

Multiplying both sides of this equality by e^{-u^2} we find

$$Je^{-u^2}=\int_{0}^{+\infty} e^{-(1+t^2)u^2}u\,dt \tag{10.52}$$

Integrating equality (10.52) with respect to u we obtain

$$J^2=J\int_{0}^{+\infty} e^{-u^2}\,du=\int_{0}^{+\infty} du\int_{0}^{+\infty} e^{-(1+t^2)u^2}u\,dt \tag{10.53}$$

Here the integrand $f(t, u) = e^{-(1+t^2)u^2}u$ is a nonnegative and continuous function for $0\leqslant t<+\infty$, $0\leqslant u<+\infty$. The inner integral in (10.53) is a continuous function of u for $0\leqslant u<+\infty$ which is implied by (10.52). If we formally reverse the order of integration we arrive at the iterated integral

$$\int_{0}^{+\infty} dt\int_{0}^{+\infty} e^{-(1+t^2)u^2}u\,du \tag{10.54}$$

whose inner integral

$$\int_{0}^{+\infty} e^{-(1+t^2)u^2}u\,du=-\frac{1}{2}\,\frac{e^{-(1+t^2)u^2}}{1+t^2}\Bigg|_{u=0}^{u=+\infty}=\frac{1}{2}\,\frac{1}{1+t^2} \tag{10.55}$$

is a continuous function of t for $0\leqslant t<+\infty$. Hence, by Theorem 10.9 on reversing the order of integration in an improper iterated integral with a nonnegative integrand, integral (10.54) is also convergent and equal to the integral (10.53). Consequently, by (10.55), we obtain

$$\int_{0}^{+\infty} dt\int_{0}^{+\infty} e^{-(1+t^2)u^2}u\,du=\frac{1}{2}\int_{0}^{+\infty}\frac{dt}{1+t^2}=\frac{\pi}{4}$$

Thus,

$$J=\int_{0}^{+\infty} e^{-x^2}\,dx=\frac{\sqrt{\pi}}{2} \tag{10.56}$$

This integral has various applications to the theory of heat conductivity, probability theory, statistical physics etc.

4. Evaluate the integral

$$J(\beta)=\int_0^{+\infty} e^{-\alpha x^2}\cos\beta x\,dx \quad \text{where} \quad \alpha=\text{const}>0 \qquad (10.57)$$

which is used in the theory of heat conductivity and statistical physics. Its convergence follows, for instance, from the fact that the integral $\int_0^{+\infty} e^{-\alpha x^2}\,dx$ is convergent. Performing formal differentiation with respect to β we obtain the equality

$$\frac{dJ}{d\beta}=\int_0^{+\infty} e^{-\alpha x^2}(-x)\sin\beta x\,dx \qquad (10.58)$$

Equality (10.58) can easily be justified. Indeed, we see that (1) the functions $e^{-\alpha x^2}\cos\beta x$ and $e^{-\alpha x^2}x\sin\beta x$ are continuous for $-\infty<\beta<+\infty$, $0\leqslant x<+\infty$, (2) integral (10.57) is convergent for $-\infty<\beta<+\infty$ and integral (10.58) converges uniformly with respect to β for $-\infty<\beta<+\infty$ by Weierstrass' test (in which we can put $g(x)=e^{-\alpha x^2}$). Therefore, by the theorem on differentiation of an improper integral with respect to the parameter, equality (10.58) is in fact valid. Now, integrating by parts in (10.58) (with respect to x) we obtain

$$\frac{dJ}{d\beta}=e^{-\alpha x^2}\frac{\sin\beta x}{2\alpha}\Bigg|_{x=0}^{x=+\infty}-\frac{\beta}{2\alpha}\int_0^{+\infty} e^{-\alpha x^2}\cos\beta x\,dx=-\frac{\beta}{2\alpha}J(\beta)$$

Separating variables in the differential equation thus obtained we find

$$\frac{dJ}{J}=-\frac{\beta\,d\beta}{2\alpha} \qquad (10.59)$$

Integrating (10.59) we obtain

$$J(\beta)=Ce^{-\frac{\beta^2}{4\alpha}}, \qquad C=\text{const} \qquad (10.60)$$

Let us determine the constant C. According to (10.56), we have

$$J(0)=\int_0^{+\infty} e^{-\alpha x^2}\,dx=\frac{1}{\sqrt{\alpha}}\int_0^{+\infty} e^{-z^2}\,dz=\frac{1}{2}\sqrt{\frac{\pi}{\alpha}} \qquad (z=x\sqrt{\alpha}) \qquad (10.61)$$

Consequently, by (10.60) and (10.61), we can write

$$J(0)=C=\frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Substituting this result into (10.60) we finally derive

$$J(\beta)=\int_0^{+\infty} e^{-\alpha x^2}\cos\beta x\,dx=\frac{1}{2}\sqrt{\frac{\pi}{\alpha}}\,e^{-\frac{\beta^2}{4\alpha}} \tag{10.62}$$

5. Let us evaluate the Fresnel integrals $\int_0^{+\infty}\sin(x^2)\,dx$ and $\int_0^{+\infty}\cos(x^2)\,dx$ which are used in optics. Putting $x^2=t$ we obtain

$$\int_0^{+\infty}\sin(x^2)\,dx=\frac{1}{2}\int_0^{+\infty}\frac{\sin t}{\sqrt{t}}\,dt,\qquad \int_0^{+\infty}\cos(x^2)\,dx=\frac{1}{2}\int_0^{+\infty}\frac{\cos t}{\sqrt{t}}\,dt$$

As an instance, let us evaluate the first integral. Noting that $\frac{1}{\sqrt{t}}=\frac{2}{\sqrt{\pi}}\int_0^{+\infty}e^{-tu^2}\,du$ (see 10.61) we derive

$$\int_0^{+\infty}\frac{\sin t\,dt}{\sqrt{t}}=\frac{2}{\sqrt{\pi}}\int_0^{+\infty}dt\int_0^{+\infty}e^{-tu^2}\sin t\,du \tag{10.63}$$

If it were easy to justify the possibility of reversing the order of integration in integral (10.63) we could complete our calculations in a simple way but it turns out that this involves some complicated techniques. Therefore, as in Example 1, we shall introduce the factor e^{-kt}, $k=\text{const}>0$, and consider the integral

$$\int_0^{+\infty}e^{-kt}\frac{\sin t}{\sqrt{t}}\,dt=\frac{2}{\sqrt{\pi}}\int_0^{+\infty}dt\int_0^{+\infty}e^{-(k+u^2)t}\sin t\,du=$$

$$=\frac{2}{\sqrt{\pi}}\int_0^{+\infty}du\int_0^{+\infty}e^{-(k+u^2)t}\sin t\,dt=\frac{2}{\sqrt{\pi}}\int_0^{+\infty}\frac{du}{1+(k+u^2)^2} \tag{10.64}$$

In the latter case it is easy to show that the order of integration can be reversed on the basis of Theorem 10.9. Since the integral $\int_0^{+\infty}e^{-kt}\frac{\sin t}{\sqrt{t}}\,dt$ converges uniformly for $0\leqslant k<+\infty$, and its integrand is continuous for $0\leqslant k<+\infty$, $0\leqslant t<+\infty$, this integral is a continuous function of k on the interval $0\leqslant k<+\infty$. Therefore, passing to the limit for $k\to 0+0$ we obtain

$$\int_0^{+\infty}\frac{\sin t}{\sqrt{t}}\,dt=\frac{2}{\sqrt{\pi}}\int_0^{+\infty}\frac{du}{1+u^4}$$

Taking the decomposition of the fraction $\frac{1}{1+u^4}$ into partial fractions and performing integration we finally obtain

$$\int_0^{+\infty} \sin(x^2)\,dx = \frac{1}{2}\sqrt{\frac{\pi}{2}} \tag{10.65}$$

The relation

$$\int_0^{+\infty} \cos(x^2)\,dx = \frac{1}{2}\sqrt{\frac{\pi}{2}} \tag{10.66}$$

is proved in a similar fashion.

6. Let us consider the **Frullani*** **integral**

$$\int_0^{+\infty} \frac{f(bx)-f(ax)}{x}\,dx \quad \text{where } 0<a<b<+\infty \tag{10.67}$$

We shall limit ourselves to the following two cases.

(1) If $f'(x)$ is continuous and integrable on the interval $0 \leqslant x < +\infty$ and $f(x)$ tends to a finite limit $f(+\infty)$ as $x \to +\infty$, i.e.

$$\int_0^{+\infty} f'(x)\,dx = f(+\infty) - f(0)$$

then the integral

$$\int_0^{+\infty} f'(ux)\,dx \tag{10.68}$$

is uniformly convergent with respect to the parameter u on the interval $0<a\leqslant u\leqslant b$. Indeed, since $f(x)$ tends to a finite limit $f(+\infty)$ as $x \to +\infty$, the necessary and sufficient condition of Cauchy's criterion for existence of a limit of a function is fulfilled for $f(x)$, that is for every $\varepsilon>0$ there is $N(\varepsilon)$ such that $|f(x')-f(x'')|<\varepsilon$ for all $x',\ x''>N(\varepsilon)$. But then we have

$$\left|\int_{A'}^{A''} f'(ux)\,dx\right| = \left|\frac{1}{u}\int_{A'u}^{A''u} f'(t)\,dt\right| = \left|\frac{f(A''u)-f(A'u)}{u}\right| \leqslant$$
$$\leqslant \frac{1}{a}\,|f(A''u)-f(A'u)| < \varepsilon$$

for all $A',\ A''>\frac{1}{a}N(\varepsilon)$ and for all u belonging to the interval $a\leqslant u\leqslant b$. Therefore integral (10.67) can be evaluated by inte-

* Frullani, Giuliano (1795-1834), an Italian mathematician.

grating (10.68) with respect to the parameter u from a to b:

$$\int_0^{+\infty} \frac{f(bx)-f(ax)}{x}\,dx = \int_0^{+\infty} dx \int_a^b f'(ux)\,du = \int_a^b du \int_0^{+\infty} f'(ux)\,dx =$$

$$= \int_a^b \frac{f(+\infty)-f(0)}{u}\,du = [f(+\infty)-f(0)] \ln \frac{b}{a} \qquad (10.69)$$

(2) If there is no finite limit of $f(x)$ for $x \to +\infty$ but the integral $\int_A^{+\infty} \frac{f(x)}{x}\,dx$, $A > 0$, is convergent, and the derivative $f'(0)$ exists, we have

$$\int_0^{+\infty} \frac{f(bx)-f(ax)}{x}\,dx = -f(0) \ln \frac{b}{a} \qquad (10.70)$$

In fact, we can write

$$\int_0^{as} \frac{f(t)-f(0)}{t}\,dt = \int_0^s \frac{f(ax)-f(0)}{x}\,dx \qquad (t=ax)$$

and

$$\int_0^{bs} \frac{f(t)-f(0)}{t}\,dt = \int_0^s \frac{f(bx)-f(0)}{x}\,dx \qquad (t=bx)$$

Consequently,

$$\int_0^s \frac{f(bx)-f(ax)}{x}\,dx = \int_{as}^{bs} \frac{f(t)}{t}\,dt - f(0) \int_{as}^{bs} \frac{dt}{t} =$$

$$= \int_{as}^{bs} \frac{f(t)}{t}\,dt - f(0) \ln \frac{b}{a}$$

whence we obtain (10.70) by passing to the limit for $s \to +\infty$.

Frullani's equalities (10.69) and (10.70) can be applied to evaluating various concrete integrals. For instance, applying (10.69), for $0 < a < b$, to the function $f(x) = e^{-x}$ we find

$$\int_0^{+\infty} \frac{e^{-bx}-e^{-ax}}{x}\,dx = \ln \frac{a}{b} \quad \text{and} \quad \int_0^{+\infty} \frac{\arctan bx - \arctan ax}{x}\,dx =$$

$$= \frac{\pi}{2} \ln \frac{b}{a}$$

Similarly, the application of (10.70) to the function $f(x) = \sin x$ results in

$$\int_0^{+\infty} \frac{\sin bx - \sin ax}{x}\, dx = 0 \quad \text{and} \quad \int_0^{+\infty} \frac{\cos bx - \cos ax}{x}\, dx = \ln \frac{a}{b}$$

$$0 < a < b$$

§ 3. EULER'S INTEGRALS

Euler's integral of the first kind

$$\mathrm{B}(p,\ q) = \int_0^1 x^{p-1}(1-x)^{q-1}\, dx$$

called the **beta function of** p **and** q **and Euler's integral of the second kind**

$$\Gamma(p) = \int_0^{+\infty} x^{p-1} e^{-x}\, dx$$

called the **gamma function of** p play an important role in various divisions of mathematics and mathematical physics. As will be shown, the beta function is expressed in terms of the gamma function (see relation (10.81)) and therefore we begin with the properties of the gamma function.

1. Properties of Gamma Function.

(1) The integral $\Gamma(p) = \int_0^{+\infty} x^{p-1} e^{-x}\, dx$ converges for $0 < p < +\infty$ and diverges for $p \leqslant 0$ (see the end of § 2, Chapter 9). This integral is improper for $p < 1$ not only because its interval of integration is infinite but also because the integrand approaches infinity for $p < 1$ as $x \to 0 + 0$.

Let us prove that the integral $\Gamma(p) = \int_0^{+\infty} x^{p-1} e^{-x}\, dx$ is uniformly convergent with respect to the parameter p on every finite interval $0 < p_0 \leqslant p \leqslant P_0 < +\infty$. As in the case of testing this integral for ordinary convergence, we break up the interval of integration $[0, +\infty)$ into two intervals, namely $0 \leqslant x \leqslant 1$ and $1 \leqslant x < +\infty$, and test, for uniform convergence, the corresponding integrals $\int_0^1 x^{p-1} e^{-x}\, dx$ and $\int_1^{+\infty} x^{p-1} e^{-x}\, dx$. The integral $\int_0^1 x^{p-1} e^{-x}\, dx$ converges uniformly for $0 < p_0 \leqslant x < +\infty$ by Weierstrass' test since $e^{-x} x^{p-1} \leqslant$

$\leqslant x^{p_0-1}$ for $0<x<1$ and $p\geqslant p_0$ and the integral $\int_0^1 x^{p_0-1}\,dx$ converges for $p_0>0$. Estimating the integral $\int_0^\lambda x^{p-1}e^{-x}\,dx$ for $p\to 0+0$ and $\lambda=\text{const}>0$ we see that

$$\int_0^\lambda x^{p-1}e^{-x}\,dx \geqslant \int_0^\lambda x^{p-1}e^{-1}\,dx = e^{-1}\int_0^\lambda x^{p-1}\,dx = \frac{\lambda^p}{pe} \to +\infty$$

and, consequently, the integral $\int_0^1 x^{p-1}e^{-x}\,dx$ does not converge uniformly on the interval $0<p<+\infty$.

Weierstrass' test indicates that the integral $\int_1^{+\infty} x^{p-1}e^{-x}\,dx$ converges uniformly for $-\infty<p\leqslant P_0<+\infty$ where P_0 is an arbitrary fixed number because we have

$$x^{p-1}e^{-x}\leqslant x^{P_0-1}e^{-x} \quad \text{for } 1\leqslant x<+\infty, \qquad -\infty<p<P_0$$

and the integral $\int_1^{+\infty} x^{P_0-1}e^{-x}\,dx$ is convergent. But the integral $\int_1^{+\infty} x^{p-1}e^{-x}\,dx$ does not converge uniformly on the interval $-\infty<p<$ $<+\infty$. To show this let us estimate the integral $\int_l^{+\infty} x^{p-1}e^{-x}\,dx$ for an arbitrary fixed $l>1$ as $p\to+\infty$. For any integer $N>0$ we have $p-1>N$, from some value of p on, when $p\to+\infty$, and therefore, for such p, we can write

$$\int_l^{+\infty} x^{p-1}e^{-x}\,dx > \int_l^{+\infty} x^N e^{-x}\,dx = -e^{-x}x^N\Big|_{x=l}^{+\infty} + N\int_l^{+\infty} x^{N-1}e^{-x}\,dx =$$

$$[l^N+Nl^{N-1}+N(N-1)\,l^{N-2}+\ldots+N!]\,e^{-l}\to+\infty$$
$$\text{for } N\to+\infty$$

Consequently,

$$\lim_{p\to+\infty}\int_l^{+\infty} x^{p-1}e^{-x}\,dx = +\infty$$

for any fixed $l>0$.

Thus, the integral $\int\limits_0^1 e^{-x}x^{p-1}\,dx$ converges uniformly on every interval $0 < p_0 \leqslant p < +\infty$ where p_0 is an arbitrary fixed positive number and the integral $\int\limits_1^{+\infty} x^{p-1}e^{-x}dx$ converges uniformly on every interval $-\infty < p \leqslant P_0 < +\infty$ where P_0 is an arbitrary finite number. Hence, both integrals simultaneously converge uniformly on every interval of the form $0 < p_0 \leqslant p \leqslant P_0 < +\infty$ and, consequently, the integral $\Gamma(p) = \int\limits_0^{+\infty} x^{p-1}e^{-x}dx$ also converges uniformly with respect to p on every such interval.

(2) Since the integrand $f(x, p) = x^{p-1}e^{-x}$ is continuous for $0 < x < +\infty$, $0 < p < +\infty$, and the integral $\int\limits_0^{+\infty} x^{p-1}e^{-x}\,dx$, understood as the limit

$$\lim_{\substack{l \to +\infty \\ \lambda \to 0+0}} \int\limits_{\lambda}^{l} x^{p-1}e^{-x}\,dx = \int\limits_0^{+\infty} x^{p-1}e^{-x}\,dx$$

is uniformly convergent with respect to p on every finite interval $0 < p_0 \leqslant p \leqslant P_0 < +\infty$, the integral $\Gamma(p) = \int\limits_0^{+\infty} x^{p-1}e^{-x}dx$ is a continuous function on every such interval, i.e. a continuous function for all p satisfying the condition $0 < p < +\infty$.

(3) Differentiating $\Gamma(p) = \int\limits_0^{+\infty} x^{p-1}e^{-x}\,dx$ formally with respect to p under the integral sign we obtain

$$\Gamma'(p) = \int\limits_0^{+\infty} x^{p-1}(\ln x)\,e^{-x}\,dx \tag{10.71}$$

But equality (10.71) can be justified because we can easily prove that the integral on the right-hand side of (10.71) is uniformly convergent on every finite interval $0 < p_0 \leqslant p \leqslant P_0 < +\infty$, and the partial derivative $f'_p(x, p) = x^{p-1}(\ln x)\,e^{-x}$ (where $f(x, p) = x^{p-1}e^{-x}$) is continuous for $0 < x < +\infty$, $0 < p < +\infty$. The fact that the integral on the right-hand side of (10.71) converges uniformly is proved by applying Weierstrass' test to the integrals

$$\int\limits_0^1 x^{p-1}(\ln x)\,e^{-x}\,dx \quad \text{and} \quad \int\limits_1^{+\infty} x^{p-1}(\ln x)\,e^{-x}\,dx$$

for which we can put, respectively, $g(x) = x^{p_0-1} \mid \ln x \mid$ and $g(x) = x^{p_0-1} \mid \ln x \mid e^{-x}$.

We likewise prove that the derivatives $\Gamma^{(k)}(p)$ of all orders $k = 1, 2, 3, \ldots$ exist and are expressed by the formulas

$$\Gamma^{(k)}(p) = \int_0^{+\infty} x^{p-1} (\ln x)^k e^{-x}\, dx, \quad k = 1, 2, \ldots \tag{10.72}$$

(4) Performing integration by parts we find

$$p\Gamma(p) = p \int_0^{+\infty} x^{p-1} e^{-x}\, dx = x^p e^{-x} \Big|_{x=0}^{+\infty} + \int_0^{+\infty} x^p e^{-x}\, dx$$

that is

$$\Gamma(p+1) = p\Gamma(p) \tag{10.73}$$

Applying recurrence formula (10.73) repeatedly we can reduce evaluation of $\Gamma(a+n)$ where $0 < a \leqslant 1$ and n is an arbitrary natural number to evaluation of $\Gamma(a)$:

$$\Gamma(a+n) = (a+n-1)(a+n-2)\ldots(a+1)\, a\Gamma(a) \tag{10.74}$$

If we put $a = 1$ and take into account that

$$\Gamma(1) = \int_0^{+\infty} e^{-x}\, dx = 1 \tag{10.75}$$

formula (10.74) results in

$$\Gamma(n+1) = n(n-1)\ldots 2\cdot 1 = n! \tag{10.76}$$

(5) Let us evaluate

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{+\infty} x^{-\frac{1}{2}} e^{-x}\, dx$$

Putting $x = t^2$ we obtain

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^{+\infty} e^{-t^2}\, dt = 2\frac{\sqrt{\pi}}{2} = \sqrt{\pi} \tag{10.77}$$

(6) Let us consider the graph of the function $\Gamma(p)$ (see Fig. 10.1). For $p \to 0+0$ and $p \to +\infty$ we have $\Gamma(p) \to +\infty$. The values of $\Gamma(p)$ for natural p are given by formula (10.76). We have $\Gamma(1) = \Gamma(2) = 1$, and therefore, by Rolle's theorem, the derivative $\Gamma'(p)$ turns into zero at a point belonging to the interval $1 < p < 2$. Let p_0 be such a point. Since $\Gamma''(p) = \int_0^{+\infty} x^{p-1} (\ln x)^2 e^{-x}\, dx > 0$ for all p, $0 < p < +\infty$, the derivative $\Gamma'(p)$ is a monotone increasing function for $0 < p < +\infty$. Consequently, the derivative $\Gamma'(p)$

has no roots on the interval $0 < p < +\infty$ other than p_0. Besides, $\Gamma'(p) < 0$ for $p < p_0$ and $\Gamma'(p) > 0$ for $p > p_0$ because $\Gamma'(p)$ is a monotone increasing function. Hence, the function $\Gamma(p)$ has

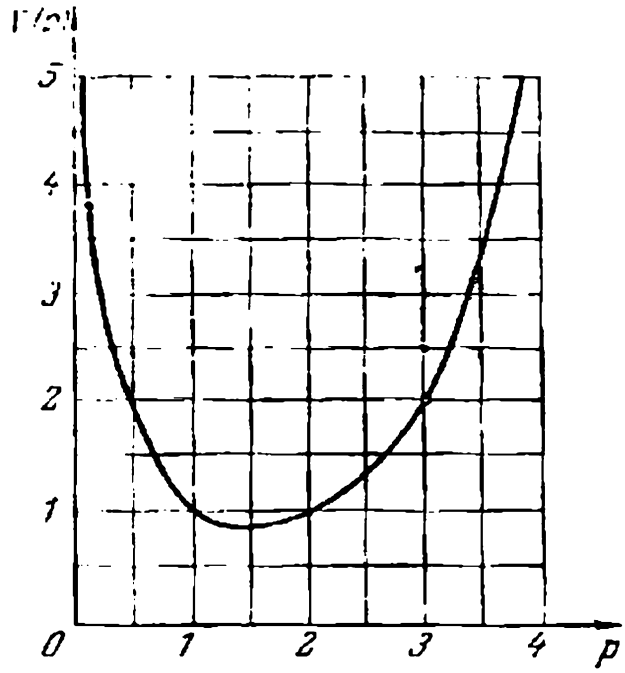


Fig. 10.1

only one extremal value on the interval $0 < p < \infty$, namely the minimum attained at the point $p = p_0$.

Let us consider the corresponding numerical data:

$$p_0 \approx 1.4616, \quad \min \Gamma(p) > \Gamma(p_0) = 0.8856$$

Since the gamma function increases for $p \geqslant 2$ we have $\Gamma(p) > \Gamma(n+1) = n!$ for $p > n+1$ where $n \geqslant 1$. Hence, $\Gamma(p) \to +\infty$ as $p \to +\infty$. Furthermore, we can write

$$\Gamma(p) = \frac{\Gamma(p+1)}{p}$$

for $p > 0$ and hence $\Gamma(p) = \dfrac{\Gamma(p+1)}{p} \to +\infty$ for $p \to 0+0$ because $\Gamma(p+1) \to \Gamma(1) = 1$ for $p \to 0+0$.

2. Properties of Beta Function.

(1) The integral $B(p, q) = \int\limits_0^1 x^{p-1}(1-x)^{q-1}\,dx$ converges for $p > 0$ and $q > 0$.

(2) The change of variable $x = 1 - t$ shows that

$$B(p, q) \equiv B(q, p) \tag{10.78}$$

Consequently, the beta function $B(p, q)$ is a symmetric function of p and q.

(3) For $q > 1$ we have

$$B(p, q) = \int_0^1 (1-x)^{q-1}\, d\left(\frac{x^p}{p}\right) =$$

$$= \frac{x^p(1-x)^{q-1}}{p}\Bigg|_{x=0}^{x=1} + \frac{q-1}{p}\int_0^1 x^p(1-x)^{q-2}\,dx =$$

$$= \frac{q-1}{p}\int_0^1 x^{p-1}(1-x)^{q-2}\,dx - \frac{q-1}{p}\int_0^1 x^{p-1}(1-x)^{q-1}\,dx =$$

$$= \frac{q-1}{p}B(p, q-1) - \frac{q-1}{p}B(p, q)$$

Thus, we have

$$B(p, q) = \frac{q-1}{p+q-1}B(p, q-1) \quad \text{for } q > 1 \tag{10.79}$$

The beta function being symmetric, relation (10.79) implies, for $p > 1$, the equality

$$B(p, q) = \frac{p-1}{p+q-1}B(p-1, q) \quad \text{for } p > 1 \tag{10.79'}$$

(4) Performing the change of variable $x = \frac{z}{1+z}$ in the integral $B(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx$ we arrive at another analytical representation of the beta function:

$$B(p, q) = \int_0^{+\infty} \frac{z^{p-1}}{(1+z)^{p+q}}\,dz \tag{10.80}$$

(5) Let us deduce a relation connecting the beta and the gamma functions. Namely, we shall prove that

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \tag{10.81}$$

for $p > 0$ and $q > 0$. The change of variable $x = tz$ ($t > 0$, $dx = t\,dz$) in the integral $\Gamma(p) = \int_0^{+\infty} x^{p-1}e^{-x}\,dx$ results in

$$\frac{\Gamma(p)}{t^p} = \int_0^{+\infty} z^{p-1}e^{-tz}\,dz \tag{10.82}$$

Substituting $1+t$ for t and $p+q$ for p we obtain from (10.82) the relation

$$\frac{\Gamma(p+q)}{(1+t)^{p+q}}=\int_0^{+\infty} z^{p+q-1}e^{-(1+t)z}\,dz \tag{10.83}$$

Multiplying both sides of the latter equality by t^{p-1} and integrating with respect to t from 0 to $+\infty$ we derive

$$\Gamma(p+q)\int_0^{+\infty}\frac{t^{p-1}}{(1+t)^{p+q}}\,dt=\int_0^{+\infty}dt\int_0^{+\infty}t^{p-1}z^{p+q-1}e^{-(1+t)z}\,dz$$

By (10.80), the last relation can be rewritten in the form

$$\Gamma(p+q)\,\mathrm{B}(p,q)=\int_0^{+\infty}dt\int_0^{+\infty}z^{p+q-1}t^{p-1}e^{-(1+t)z}\,dz \tag{10.84}$$

Now let us prove that it is allowable to reverse the order of integration in integral (10.84) for $p>1$ and $q>1$. Indeed, we can easily show that the conditions of Theorem 10.9 on reversing the order of integration in an improper iterated integral are fulfilled here:

(a) the function

$$f(z,t)=z^{p+q-1}t^{p-1}e^{-(1+t)z}\geqslant 0$$

is continuous for $0\leqslant z<+\infty$, $0\leqslant t<+\infty$;

(b) if $p>1$ and $q>1$ integral (10.84) is convergent;

(c) the integral

$$\int_0^{+\infty}t^{p-1}z^{p+q-1}e^{-(1+t)z}\,dt=\Gamma(p+q)\frac{t^{p-1}}{(1+t)^{p+q}}$$

is a continuous function of t for $0\leqslant t<+\infty$ and the integral

$$\int_0^{+\infty}t^{p-1}z^{p+q-1}e^{-(1+t)z}\,dt=\Gamma(p)\,z^{q-1}e^{-z}$$

is a continuous function of z for $0\leqslant z<+\infty$.

Hence, the iterated integral $\int_0^{+\infty}dz\int_0^{+\infty}z^{p+q-1}t^{p-1}e^{-(1+t)z}\,dt$ is convergent and equal to integral (10.84) (Theorem 10.9).

Consequently,

$$\Gamma(p+q)\,B(p,\ q)=\int_0^{+\infty} dz \int_0^{+\infty} z^{p+q-1}t^{p-1}e^{-(1+t)z}\,dt=$$

$$=\int_0^{+\infty} z^{p+q-1}e^{-z}\,dz \int_0^{+\infty} t^{p-1}e^{-tz}\,dt=\int_0^{+\infty} z^{p+q-1}e^{-z}\frac{\Gamma(p)}{z^p}\,dz=$$

$$=\Gamma(p)\int_0^{+\infty} z^{q-1}e^{-z}\,dz=\Gamma(p)\,\Gamma(q)$$

We have thus proved that equality (10.81) is valid for $p>1$ and $q>1$. To extend relation (10.81) to all $p>0$ and $q>0$ we simply write this relation for $p>1$ and $q>1$ and then apply recurrence formulas (10.79) and (10.79′) to its left-hand side and recurrence formula (10.73) to its right-hand side.

(6) Let us derive the formula

$$B(p,\ 1-q)=\frac{\pi}{\sin p\pi}\quad \text{for } 0<p<1 \tag{10.85}$$

Substituting $q=1-p$ into formula (10.80) we obtain

$$B(p,\ 1-q)=\int_0^{+\infty}\frac{z^{p-1}}{1+z}\,dz,\quad 0<p<1 \tag{10.86}$$

But in Sec. 5 of § 2 (Example 1) we showed that integral (10.86) is equal to $\frac{p}{\sin p\pi}$ for $0<p<1$, and this implies relation (10.85).

Making use of formulas (10.81) and (10.85) we obtain the formula

$$\Gamma(p)\,\Gamma(1-p)=\frac{\pi}{\sin p\pi}\quad \text{for } 0<p<1 \tag{10.87}$$

which plays an important role in the theory of the gamma function.

There are many integrals that we can evaluate by reducing them to Euler's integrals.

Examples

1. $$\int_0^{+\infty}\frac{\sqrt[4]{x}}{(1+x)^2}\,dx=B\left(\frac{5}{4},\ \frac{3}{4}\right)=\frac{\Gamma\left(\frac{5}{4}\right)\Gamma\left(\frac{3}{4}\right)}{\Gamma(2)}=$$

$$=\frac{1}{4}\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right)=\frac{1}{4}\,\frac{\pi}{\sin\frac{\pi}{4}}=\frac{\pi\sqrt{2}}{4}$$

2. Let us evaluate the integral $J=\int\limits_0^{\frac{\pi}{2}}\sin^{\frac{5}{2}}x\cos^{\frac{3}{2}}x\,dx$. Putting $\sin^2 x=z$ we obtain $\sin x=z^{\frac{1}{2}}$, $\cos x=(1-z)^{\frac{1}{2}}$ and $dz=2\sin x\cos x\,dx$. Consequently, taking into account the foregoing example and formula (10.79′) we find

$$J=2\mathrm{B}\left(\frac{7}{4},\ \frac{5}{4}\right)=\frac{3\pi\sqrt{2}}{16}$$

3. Let us consider the integral

$$J=\int\limits_0^{\frac{\pi}{2}}\sin^{p-1}x\cos^{q-1}x\,dx,\quad p>0,\quad q>0$$

Putting $\sin^2 x=z$ we get

$$\int\limits_0^{\frac{\pi}{2}}\sin^{p-1}x\cos^{q-1}x\,dx=\frac{1}{2}\int\limits_0^1 z^{\frac{p}{2}-1}(1-z)^{\frac{q}{2}-1}\,dz=$$

$$=\frac{1}{2}\,\mathrm{B}\left(\frac{p}{2},\ \frac{q}{2}\right)=\frac{1}{2}\,\frac{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)}{\Gamma\left(\frac{p+q}{2}\right)}$$

In particular, for $q=1$ we obtain the formula

$$\int\limits_0^{\frac{\pi}{2}}\sin^{p-1}x\,dx=\frac{\sqrt{\pi}}{2}\,\frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{p+1}{2}\right)}$$

§ 4. MULTIPLE INTEGRALS DEPENDENT ON PARAMETER

For definiteness, we shall consider triple integrals dependent on a parameter although the results obtained below hold for multiple integrals of an arbitrary order except certain cases which will be stipulated in what follows.

Let a function $f(x,\ y,\ z,\ \alpha,\ \beta,\ \gamma)$ be defined for $(x,\ y,\ z)\in\Omega$ and $(\alpha,\ \beta,\ \gamma)\in\Omega^*$ where Ω and Ω^* are, respectively, domains of the x, y, z-space and the α, β, γ-space. Suppose that the integral

$$J(\alpha,\ \beta,\ \gamma)=\iiint\limits_{\Omega} f(x,\ y,\ z,\ \alpha,\ \beta,\ \gamma)\,dx\,dy\,dz \qquad (10.88)$$

exists as a proper or improper integral for all $(\alpha, \beta, \gamma) \in \Omega^*$. Then it is a function of the parameters α, β, γ in the domain Ω^*.

As in the case of a onefold integral, we can easily prove the following assertions:

(1) *If $f(x, y, z, \alpha, \beta, \gamma)$ is continuous as a function defined in the domain $\overline{\Omega} \times \overline{\Omega}^*$ where $\overline{\Omega}$ and $\overline{\Omega}^*$ are bounded closed domains the integral $J(\alpha, \beta, \gamma)$ is a continuous function of the parameters α, β, γ in the domain $\overline{\Omega}^*$.**

(2) *If, in addition, the derivative $f'_\alpha(x, y, z, \alpha, \beta, \gamma)$ exists and is continuous in $\overline{\Omega} \times \overline{\Omega}^*$, the integral $J(\alpha, \beta, \gamma)$ can be differentiated according to the rule*

$$\frac{\partial J}{\partial \alpha} = \frac{\partial}{\partial \alpha} \iiint_{\Omega} f(x, y, z, \alpha, \beta, \gamma)\, dx\, dy\, dz =$$

$$= \iiint_{\Omega} f'_\alpha(x, y, z, \alpha, \beta, \gamma)\, dx\, dy\, dz \qquad (10.89)$$

The derivatives of J with respect to β and γ are expressed similarly provided that the derivatives f'_β and f'_γ exist and are continuous.

(3) *If the conditions of Assertion* (1) *are fulfilled, it is permissible to integrate the integral $J(\alpha, \beta, \gamma)$ with respect to the parameters α, β and γ under the integral sign.*

Assertions (1)-(3) are proved by analogy with the case of a single integral. They can also be easily extended to integrals of the form

$$J(\alpha, \beta, \gamma) = \iiint_{\Omega} f(x, y, z, \alpha, \beta, \gamma)\, g(x, y, z)\, dx\, dy\, dz \qquad (10.90)$$

where $f(x, y, z, \alpha, \beta, \gamma)$ satisfies the above requirements and

$$\iiint_{\Omega} |g(x, y, z)|\, dx\, dy\, dz < K = \text{const} < +\infty$$

(here the integral $\iiint_{\Omega} |g(x, y, z)|\, dx\, dy\, dz$ may be proper or improper).

Example. The potential function of the field of gravitation generated by a material body Ω with volume mass density $\rho(M) = \rho(x, y, z)$ is expressed, at every point $O(x_0, y_0, z_0)$ lying outside

---

* For given domains $\overline{\Omega}$ and $\overline{\Omega}^*$ ($(x, y, z) \in \overline{\Omega}$, $(\alpha, \beta, \gamma) \in \overline{\Omega}^*$), the domain $\overline{\Omega} \times \overline{\Omega}^*$ is the set of all points $(x, y, z, \alpha, \beta, \gamma)$ (belonging to the corresponding 6-dimensional Euclidean space) obtained when the points (x, y, z) and (α, β, γ) independently run through their domains $\overline{\Omega}$ and $\overline{\Omega}^*$.

the body, by the integral

$$U(Q) = U(x_0, y_0, z_0) = \iiint_\Omega \frac{\rho(P)}{r_{PQ}}\, dx\, dy\, dz \qquad (10.91)$$

where $r_{PQ} = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$ is the distance between the points $P(x, y, z)$ and $Q(x_0, y_0, z_0)$. If the point Q is at a positive distance from the body Ω the function $f(x, y, z, x_0, y_0, z_0) = \frac{1}{r_{PQ}}$ is continuous and possesses the continuous partial derivatives $\frac{\partial f}{\partial x_0}$, $\frac{\partial f}{\partial y_0}$ and $\frac{\partial f}{\partial z_0}$. The density $\rho(x, y, z)$ is supposed to be an absolutely integrable function in Ω. Differentiating (10.91) with respect to x_0, y_0 and z_0 according to Leibniz' rule (see relation (10.89)) we find the projections on the coordinate axes of the force of attraction with which the body Ω acts upon unit mass located at the point $Q(x_0, y_0, z_0)$:

$$\left.\begin{aligned} F_x(Q) &= \frac{\partial U}{\partial x_0} = \iiint_\Omega \frac{\rho(P)}{r^3_{PQ}}(x-x_0)\, dx\, dy\, dz \\ F_y(Q) &= \frac{\partial U}{\partial y_0} = \iiint_\Omega \frac{\rho(P)}{r^3_{PQ}}(y-y_0)\, dx\, dy\, dz \\ F_z(Q) &= \frac{\partial U}{\partial z_0} = \iiint_\Omega \frac{\rho(P)}{r^3_{PQ}}(z-z_0)\, dx\, dy\, dz \end{aligned}\right\} \qquad (10.92)$$

If the point $Q(x_0, y_0, z_0)$ lies inside the body Ω we have $r_{PQ} = 0$ when P coincides with Q. Hence, in this case Q is a singular point of the integrand in the integrals (10.91) and (10.92) and thus these integrals are improper even if $\rho(P) = \rho(x, y, z)$ is a bounded integrable function in Ω. A feature of these improper integrals dependent on the parameters x_0, y_0, z_0 (on the point $Q(x_0, y_0, z_0)$) is that the coordinates of the singular point of the integrand depend on these parameters, namely, are equal to them. Here we shall limit ourselves to studying improper integrals dependent on parameters which are of the form

$$J(Q) = \iiint_\Omega F(P, Q)\, f(P)\, dx\, dy\, dz \qquad (10.93)$$

where

$$P(x, y, z) \in \Omega \quad \text{and} \quad Q(x_0, y_0, z_0) \in \Omega$$

The function $F(P, Q)$ is supposed to be continuous for $P \neq Q$ and unbounded for $P \to Q$ while $f(P)$ is a bounded integrable function on Ω (integrals (10.91) and (10.92) are special cases of integral (10.93)).

Definition. *We say that integral* (10.93) *is* ***uniformly convergent at a point*** $Q(x_0, y_0, z_0) \in \Omega$ *if for every* $\varepsilon > 0$ *there is* $\delta = \delta(\varepsilon) > 0$ *such that the inequality*

$$\left| \iiint\limits_{\Omega_{\delta(\varepsilon)}} F(P, Q') f(P)\, dx\, dy\, dz \right| < \varepsilon \tag{10.94}$$

is fulfilled for any domain $\Omega_{\delta(\varepsilon)}$ *of diameter less than* $\delta(\varepsilon)$, *belonging to* Ω *and containing the point* Q, *and for any point* Q' *whose distance from* Q *is less than* $\delta(\varepsilon)$.

Sufficient Condition for Uniform Convergence. *If there exist a neighbourhood of the point* $Q(x_0, y_0, z_0) \in \Omega$ *and constants* $C > 0$ *and* $\lambda < 3$ *such that for all* P *and* Q' *belonging to this neighbourhood the inequality*

$$|F(P, Q')| \leqslant \frac{C}{r_{PQ'}^{\lambda}} \quad (\lambda = \text{const} < 3,\ 0 < C = \text{const} < +\infty) \tag{10.95}$$

holds, integral (10.93) *is uniformly convergent at the point* $Q(x_0, y_0, z_0)$.

Proof. By the hypothesis, we have $|f(P)| < K = \text{const} < +\infty$ everywhere in Ω. Consequently, if a ball $S_{\delta(\varepsilon)}(Q)$ of radius $\delta(\varepsilon)$ with centre at Q lies within the above neighbourhood of the point Q, we can write the inequality

$$\left| \iiint\limits_{\Omega_{\delta(\varepsilon)}} F(P, Q') f(P)\, dx\, dy\, dz \right| \leqslant$$

$$\leqslant \iiint\limits_{\Omega_{\delta(\varepsilon)}} \frac{C}{r_{PQ'}^{\lambda}} K\, dx\, dy\, dz \leqslant CK \iiint\limits_{S_{2\delta(\varepsilon)}(Q')} \frac{dx\, dy\, dz}{r_{PQ'}^{\lambda}} \tag{10.96}$$

for every domain $\Omega_{\delta(\varepsilon)}$ of diameter less than $\delta(\varepsilon)$ containing the point Q and for any point $Q' \in S_{\delta(\varepsilon)}$ where $S_{2\delta(\varepsilon)}(Q')$ is a ball of radius $2\delta(\varepsilon)$ with centre at Q' (see Fig. 10.2). Passing to spherical coordinates with origin at the point Q' we obtain

$$\iiint\limits_{S_{2\delta(\varepsilon)}} \frac{dx\, dy\, dz}{r_{PQ'}^{\lambda}} = \int\limits_0^{2\pi} d\varphi \int\limits_0^{\pi} d\theta \int\limits_0^{2\delta(\varepsilon)} \frac{r^2 \sin\theta}{r^{\lambda}}\, dr =$$

$$= 4\pi \int\limits_0^{2\delta(\varepsilon)} r^{2-\lambda}\, dr = \frac{4\pi}{3-\lambda} [2\delta(\varepsilon)]^{3-\lambda} \tag{10.96$'$}$$

It follows from (10.96) and (10.96′) that

$$\left| \iiint\limits_{\Omega_{\delta(\varepsilon)}} F(P, Q') f(P)\, dx\, dy\, dz \right| \leqslant \frac{4\pi}{3-\lambda} [2\delta(\varepsilon)]^{3-\lambda} CK \tag{10.97}$$

Since $3 - \lambda > 0$, the right-hand side of (10.97) is less than ε if $\delta(\varepsilon)$ is sufficiently small, which is what we set out to prove.

Note. In the case of an improper N-fold multiple integral of type (10.93) ($N \geqslant 1$) the exponent λ must satisfy the inequality $\lambda < N$.

If mass density ρ (P) in integrals (10.91) and (10.92) is a bounded integrable function in Ω the above test implies that these integrals are uniformly convergent at every point Q $(x_0, y_0, z_0) \in \Omega$.

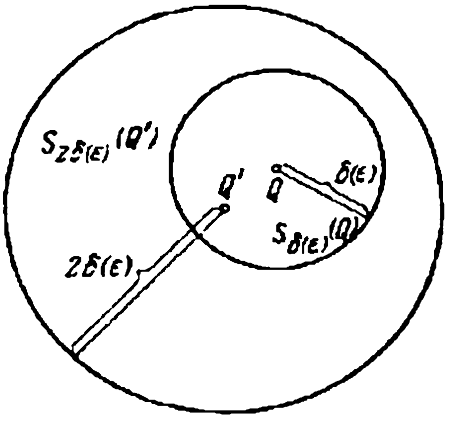


Fig. 10.2

The implications of uniform convergence are the same as in the case of onefold integrals. As an instance, let us consider the question of continuity and differentiability of an integral as a function of its parameters.

Theorem 10.10. *If integral* (10.93) *is uniformly convergent at a point $Q \in \Omega$ and the functions F (P, Q) and f (P) satisfy the above conditions, integral* (10.93) *is a continuous function at the point Q.*

Proof. Let us show that for every $\varepsilon > 0$ there is $\delta = \delta\ (\varepsilon) > 0$ such that the inequality $|\, r_{QQ'}\, | < \delta\ (\varepsilon)$ implies the inequality

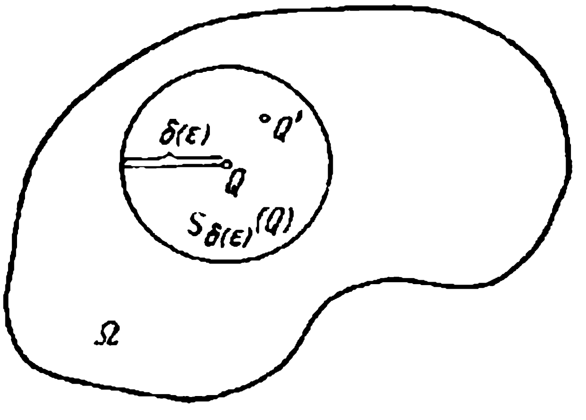


Fig. 10.3

$|\, J\ (Q) - J\ (Q')\, | < \varepsilon$. For this purpose we take a ball $S_{\delta(\varepsilon)}\ (Q)$ of a sufficiently small radius $\delta\ (\varepsilon)$ with centre at Q lying inside Ω (see Fig. 10.3) and represent each of the integrals $J\ (Q)$ and $J\ (Q')$ as a sum of two terms, namely an integral J_1 over the domain $S_{\delta(\varepsilon)}\ (Q)$ and another integral J_2 over the domain $\Omega - S_{\delta(\varepsilon)}\ (Q)$. Then we have

$$|\, J\ (Q) - J\ (Q')\, | \leqslant |\, J_2\ (Q) - J_2\ (Q')\, | + |\, J_1\ (Q) | + | J_1\ (Q') | \tag{10.98}$$

If $\delta(\varepsilon) > 0$ is sufficiently small the second and the third terms on the right-hand side of (10.98) are less than $\frac{\varepsilon}{3}$ because the integral J_1 is uniformly convergent at the point Q. Let us choose an arbitrary positive $\delta'(\varepsilon) < \frac{1}{2}\delta(\varepsilon)$. Then, if the distance between Q and Q' satisfies the inequality

$$|\overline{QQ'}| < \delta'(\varepsilon) \tag{10.99}$$

the integrals in the first summand on the right-hand side of (10.98) are proper. Consequently, if $\delta'(\varepsilon)$ $(0 < \delta'(\varepsilon) < \frac{1}{2}\delta(\varepsilon))$ is sufficiently small the theorem on the continuous dependence of a proper integral on parameters implies that the first summand on the right-hand side of (10.98) is also less than $\frac{\varepsilon}{3}$ provided that condition (10.99) is fulfilled. These results indicate that (10.98) and (10.99) imply the inequality

$$|J(Q) - J(Q')| < \varepsilon \tag{10.100}$$

which is what we set out to prove.

Here we do not discuss the general problem of differentiability of improper integrals of form (10.93) with respect to parameters and refer the reader to special courses (e.g. see [15], Lecture 7, § 2). We shall only illustrate the corresponding techniques in connection with a special case when $F(P, Q) = \frac{1}{r_{PQ}}$ and $f(P) = \rho(P)$ where $\rho(P)$ is a bounded $(|\rho(P)| < C = \text{const} < +\infty)$ and differentiable function in Ω, that is when the integral in question is of form (10.91). This case is important for the theory of potential functions (see [17]).

If the point $Q(x_0, y_0, z_0)$ lies outside Ω (and $P(x, y, z)$ runs through Ω) integral (10.91) is proper and, as has been shown, equalities (10.92) are valid. Let us prove that relations (10.92) remain valid when the point $Q(x_0, y_0, z_0)$ belongs to Ω. We shall limit ourselves to the first equality (10.92). To prove this equality we must show that the difference

$$\alpha = \frac{U(x_0+\Delta x_0, y_0, z_0) - U(x_0, y_0, z_0)}{\Delta x_0} - \\ - \iiint_{\Omega} \frac{\rho(P)(x - x_0)}{r_{PQ}^3}\, dx\, dy\, dz, \quad \Delta x_0 \neq 0 \tag{10.101}$$

tends to zero for every fixed point $Q(x_0, y_0, z_0)$ belonging to Ω as $\Delta x_0 \to 0$. Let there be given an arbitrary $\varepsilon > 0$. Take a ball $S_{\delta(\varepsilon)}(Q)$ of a sufficiently small radius $\delta(\varepsilon)$ lying entirely inside Ω and denote, respectively, by $U_1(x_0, y_0, z_0)$ and $U_2(x_0, y_0, z_0)$ the

integrals of type (10.91) over the domains $\Omega_1 = S_{\delta(\varepsilon)}(Q)$ and $\Omega_2 = \Omega - \Omega_1 = \Omega - S_{\delta(\varepsilon)}(Q)$. We obviously have $U = U_1 + U_2$ and therefore difference (10.101) can be rewritten in the form

$$\alpha = \left\{\frac{\Delta U_1}{\Delta x_0}\right\} + \left\{-\iiint\limits_{\Omega_1} \frac{\rho(P)(x-x_0)}{r_{PQ}^3}\,dx\,dy\,dz\right\} + \left\{\frac{\Delta U_2}{\Delta x_0} - \iiint\limits_{\Omega_2} \frac{\rho(P)(x-x_0)}{r_{PQ}^3}\,dx\,dy\,dz\right\} \qquad (10.102)$$

Let us estimate the first term on the right-hand side of (10.102). We have

$$\frac{\Delta U_1}{\Delta x_0} = \frac{1}{\Delta x_0}\iiint\limits_{\Omega_1} \rho(P)\left(\frac{1}{r_{PQ'}} - \frac{1}{r_{PQ}}\right)dx\,dy\,dz = -\frac{1}{\Delta x_0}\iiint\limits_{\Omega_1} \rho(P)\,\frac{r_{PQ} - r_{PQ'}}{r_{PQ}r_{PQ'}}\,dx\,dy\,dz \qquad (10.103)$$

where the point Q' belongs to $\Omega_1 = S_{\delta(\varepsilon)}(Q)$ and has the coordinates $(x_0 + \Delta x_0, y_0, z_0)$. The sides of the triangle QPQ' are of lengths r_{PQ}, $r_{PQ'}$ and $|\Delta x_0|$ and therefore

$$|r_{PQ'} - r_{PQ}| \leqslant |\Delta x_0| \qquad (10.104)$$

Taking into account inequality (10.104) and the apparent relation

$$\frac{1}{r_{PQ}r_{PQ'}} \leqslant \frac{1}{2}\left(\frac{1}{r_{PQ}^2} + \frac{1}{r_{PQ'}^2}\right)$$

we derive the following estimate for (10.103):

$$\left|\frac{\Delta U_1}{\Delta x_0}\right| \leqslant \frac{C}{2}\iiint\limits_{\Omega_1}\left(\frac{1}{r_{PQ}^2} + \frac{1}{r_{PQ'}^2}\right)dx\,dy\,dz \qquad (10.105)$$

since $|\rho(P)| \leqslant C$. The integral $\iiint\limits_{\Omega_1} \frac{dx\,dy\,dz}{r_{PQ}^2}$ converges uniformly on $\Omega_1 = S_{\delta(\varepsilon)}(Q)$ (see the sufficient condition for uniform convergence) and, consequently, the right-hand side of inequality (10.105) is less than $\frac{\varepsilon}{3}$ if $\delta = \delta(\varepsilon) > 0$ is sufficiently small. The second term on the right-hand side of (10.102) is a uniformly convergent improper integral, which is implied by the inequality $\frac{|x - x_0|}{r_{PQ}} \leqslant 1$ and the sufficient condition for uniform convergence. Hence, this term is less than $\frac{\varepsilon}{3}$ in its modulus for all sufficiently small $\delta = \delta(\varepsilon) > 0$. Finally, let us estimate the third term on the right-hand side of (10.102). Since the integral $U_2(x_0, y_0, z_0) = \iiint\limits_{\Omega_2} \frac{\rho(P)}{r_{PQ}}\,dx\,dy\,dz$ is of form (10.90) (because the point

$Q(x_0, y_0, z_0)$ lies outside Ω_2) it can be differentiated with respect to the parameter x_0 under the sign of integration and, consequently, for all sufficiently small $|\Delta x_0| < \delta = \delta(\varepsilon)$ we have

$$\left| \frac{\Delta U_2}{\Delta x_0} - \iiint\limits_{\Omega_2} \frac{\rho(P)(x - x_0)}{r_{PQ}^3} \, dx \, dy \, dz \right| < \frac{\varepsilon}{3} \tag{10.106}$$

Thus, for all sufficiently small $\delta = \delta(\varepsilon)$ and $|\Delta x_0|$, difference (10.101) is less than ε, which is what we set out to prove.

Note. The results obtained here for volume integrals can easily be extended to line integrals and surface integrals. Let the reader modify appropriately the formulations and proofs of the above theorems.

11 Fourier Series and Fourier Integral

In natural and engineering sciences we often deal with *periodic processes* such as oscillatory or rotary motion of various parts of machines and apparatuses, periodic motion of heavenly bodies and elementary particles, acoustic and electromagnetic vibrations etc.

These processes are described mathematically by means of periodic functions. *A function $f(t)$ of one independent variable t is said to be* **periodic** *if there is a number $T \neq 0$ (called its* **period**) *such that*

$$f(t+T) \equiv f(t) \quad \textit{for all values of } t, \quad -\infty < t < +\infty \tag{11.1}$$

The well known trigonometric functions $\sin t$ and $\cos t$ with period $T = 2\pi$ are the simplest periodic functions.

In physics the simplest periodic processes are described by means of the function

$$\xi(t) = A \sin(\omega t + \varphi), \quad -\infty < t < +\infty \tag{11.2}$$

called a **harmonic** (or a **harmonic vibration**). We have

$$\xi\left(t + \frac{2\pi}{\omega}\right) \equiv \xi(t) \quad \text{for } -\infty < t < +\infty \tag{11.3}$$

and therefore the quantity $T = \frac{2\pi}{\omega}$ is a period of this harmonic. The constants A and ω are called, respectively, the **amplitude** and the **frequency** of the harmonic, the expression $\omega t + \varphi$ is called its **phase** and the constant φ is the **initial phase**.

The fundamental question this chapter is devoted to concerns the problem of representing an arbitrary periodic function in the form of a sum of harmonics.

§ 1. PROPERTIES OF PERIODIC FUNCTIONS. STATEMENT OF THE KEY PROBLEM

1. Periods of a Periodic Function. Let $f(t)$ be a periodic function with period $T \neq 0$, that is

$$f(t+T) \equiv f(t) \quad \text{for all } t, \quad -\infty < t < +\infty \tag{11.4}$$

Then every integer multiple kT, $k = \pm 1, \pm 2, \ldots$, of the period T also serves as a period of this function.

Indeed, if T is a period we have, for any integer $k > 1$, the relation

$$f(t + kT) = f[t + (k-1)T + T] = \\ = f[t + (k-1)T] = \ldots = f(t) \qquad (11.5)$$

for all t, $-\infty < t < +\infty$, which means that kT is a period of $f(t)$. Furthermore, we have

$$f(t - T) = f[(t - T) + T] = f(t) \quad \text{for all } t, \\ -\infty < t < +\infty \qquad (11.6)$$

and consequently the number $-T$ is a period of $f(t)$. But then it follows from (11.5) that the number $k(-T) = -kT$ is also a period of $f(t)$ for any integer $k > 1$. Thus, the above assertion has been proved.

We can now easily verify that if two numbers T_1 and T_2 are periods of a function $f(t)$ the numbers $T_1 \pm T_2$ are also periods of this function.

A constant quantity can obviously be regarded as a periodic function with an arbitrary period, i.e. every number is its period.

*If $f(t)$ is a continuous periodic function which is not identically equal to a constant it possesses the least positive period called its **primitive period**.** When speaking about the period of a function we usually mean its primitive period.

2. Periodic Extension of a Nonperiodic Function. Taking an arbitrary nonperiodic function $f(x)$** defined on an interval $a \leqslant x \leqslant a + T$ we can construct a periodic function $F(x)$ with period T which coincides with $f(x)$ on this interval. Geometrically, to obtain the graph of $F(x)$ we must translate the graph of the function $f(x)$ (constructed on the interval $a \leqslant x \leqslant a + T$) along the x-axis to the left and right by shifting this graph, in succession, by distances $T, 2T, 3T, \ldots, nT, \ldots$ as is shown in Fig. 11.1. This process (and the resulting function $F(x)$ itself) is called the

---

* If we suppose that a continuous periodic function $f(t)$ different from a constant does not have the least positive period, it can easily be shown that there must exist a sequence of its positive periods $T_1, T_2, \ldots, T_n, \ldots$ convergent to zero. The set of all the integer multiples of these periods constitute an everywhere dense point set lying on the t-axis, $-\infty < t < +\infty$, and, consequently, the values of $f(t)$ assumed at the points of this everywhere dense set are equal to its value at the origin of coordinates. Therefore $f(t)$ is identically equal to the constant $f(0)$ on this set and hence, by the continuity of the function, it must be identically equal to this constant on the whole t-axis, $-\infty < t < +\infty$, which contradicts the hypothesis. Hence, the smallest positive period must exist.

** In what follows we shall denote the independent variable by [illegible]

periodic extension, with period T, of the function $f(x)$ from the interval $a \leqslant x \leqslant a + T$ to the x-axis. In the general case this

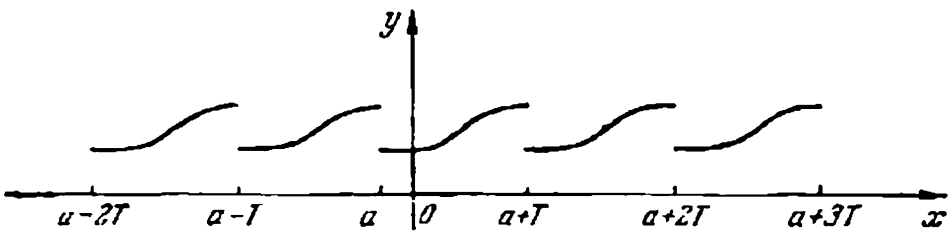


Fig. 11.1

process does not uniquely specify the values of $F(x)$ at the points $x = a \pm kT$, $k = 1, 2, 3, \ldots$.

3. Integral of a Periodic Function. If $f(x)$ is a periodic integrable function with period T we have $\int\limits_a^{a+T} f(x)\,dx \equiv \int\limits_0^T f(x)\,dx$ for any a, $-\infty < a < +\infty$. In fact, we can write

$$\int\limits_a^{a+T} f(x)\,dx = \int\limits_a^T f(x)\,dx + \int\limits_T^{a+T} f(x)\,dx = \int\limits_a^T f(x)\,dx + \int\limits_0^a f(x)\,dx = \int\limits_0^T f(x)\,dx$$

because the periodicity of the function implies that

$$\int\limits_T^{a+T} f(x)\,dx = \int\limits_T^{a+T} f(x-T)\,dx = \int\limits_0^a f(x')\,dx' \quad \text{where } x' = x - T$$

Thus, the integral of a periodic function with period T taken over an arbitrary interval of length T has one and the same value.

4. Arithmetical Operations on Periodic Functions. It is apparent that a sum, a difference, a product and a ratio of periodic functions with the same period T are again periodic functions with period T.

Let $f(x)$ and $g(x)$ be two periodic functions. If their periods T_f and T_g are *commensurable*, that is $\frac{T_f}{T_g} = \frac{p}{q}$ where p and q are integers, the number $T^* = pT_g = qT_f$ is a period of both functions $f(x)$ and $g(x)$. Consequently, the sum, the difference, the product and the ratio of these functions are also periodic functions with period T^*.

But if the periods T_f and T_g of the functions $f(x)$ and $g(x)$ are *incommensurable* their sum is no longer a periodic function but a so-called *almost periodic function*.

Let us give the definition of an almost periodic function. *A function $f(x)$ continuous on the entire real axis $-\infty < x < +\infty$ is said to be* **almost periodic** *if for every $\varepsilon > 0$ there exists a number $L = L(\varepsilon) > 0$ such that for every interval of length L, $\alpha \leqslant x \leqslant \alpha + L$, $-\infty < \alpha < +\infty$, there is at least one number $\tau = \tau(\varepsilon)$*

such that

$$| f (x + \tau (\varepsilon)) - f (x) | < \varepsilon$$

for all x, $-\infty < x < +\infty$.

Periodic functions are obviously a special case of almost periodic functions. It can be proved that a sum, a difference, a product and a ratio of almost periodic functions (under the condition that in the latter the divisor is different from zero) are almost periodic functions. Hence, the set of all almost periodic functions (in contrast to the set of all periodic functions) is closed with respect to the fundamental operations of arithmetic.

5. Superposition of Harmonics with Multiple Frequencies. Let us consider a sequence of harmonics

$$A_k \sin \left(\frac{2\pi k}{T} x + \varphi_k \right), \quad k = 1, 2, \ldots, \quad -\infty < x < +\infty, \quad T > 0 \tag{11.7}$$

Obviously, the number $T_k = \frac{T}{k}$ is the period of the kth harmonic.* Consequently, the number $T = kT_k$ is a common period for all the harmonics entering into sequence (11.7) $\Big($but it is the least period only for the first harmonic $A_1 \sin \left(\frac{2\pi}{T} x + \varphi_1 \right) \Big)$. The frequency of the kth harmonic is equal to $\lambda_k = \frac{2\pi k}{T}$, $k = 1, 2, \ldots$. Thus, the frequencies of the harmonics belonging to sequence (11.7) are integer multiples of one and the same number $\frac{2\pi}{T}$. The least positive (primitive) period of the kth harmonic is equal to $T_k = \frac{T}{k}$, $k = 1, 2, \ldots$, and, consequently, we have $\lambda_k = \frac{2\pi}{T_k}$ and $\lambda_k = k\lambda_1$. Such harmonics will be referred to as the ones with *multiple frequencies*.

A sum (or, in physical terminology, a superposition) of a finite number of harmonics of the form

$$f_N (x) = A_0 + \sum_{k=1}^{N} A_k \sin \left(\frac{2\pi k}{T} x + \varphi_k \right) \tag{11.8}$$

---

* In fact, we have $\sin \left[\frac{2\pi k}{T} \left(x + \frac{T}{k} \right) + \varphi_k \right] = \sin \left[\left(\frac{2\pi k}{T} x + \varphi_k \right) + 2\pi \right] = \sin \left(\frac{2\pi k}{T} x + \varphi_k \right)$.

is a periodic function with period T since the number T is a period of all the harmonics* and the least period of the first harmonic.

Similarly, a superposition of an infinite number of such harmonics, or, more precisely, the sum of a convergent series of the form

$$f(x) = A_0 + \sum_{k=1}^{+\infty} A_k \sin\left(\frac{2\pi k}{T} x + \varphi_k\right) \tag{11.9}$$

is also a periodic function with period T.

Equalities (11.8) and (11.9) can be transformed in the following way. Taking into account that

$$A_k \sin\left(\frac{2\pi k}{T} x + \varphi_k\right) = A_k \sin\varphi_k \cos\frac{2\pi k}{T} x + A_k \cos\varphi_k \sin\frac{2\pi k}{T} x$$

we put

$$\frac{a_0}{2} = A_0, \quad a_k = A_k \sin\varphi_k, \quad b_k = A_k \cos\varphi_k, \quad 2l = T$$

and rewrite (11.8) and (11.9) in the form

$$f_N(x) = \frac{a_0}{2} + \sum_{k=1}^{N} \left(a_k \cos\frac{k\pi x}{l} + b_k \sin\frac{k\pi x}{l}\right) \tag{11.10}$$

and

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{+\infty} \left(a_k \cos\frac{k\pi x}{l} + b_k \sin\frac{k\pi x}{l}\right) \tag{11.11}$$

All the functions on the right- and left-hand sides of (11.10) and (11.11) are periodic, with period $2l$.

It should be noted that the functions $f_N(x)$ and $f(x)$ are of a more complicated nature in comparison with the harmonics $\cos\frac{k\pi x}{l}$ and $\sin\frac{k\pi x}{l}$, $k = 1, 2, \ldots$, they are formed of (e.g. see Fig. 11.7b).

The series on the right-hand side of (11.11) *is called a* ***trigonometric series***. *A relation of form* (11.11) (*provided it is valid*) *is called the* ***expansion of the function*** $f(x)$ ***into a trigonometric series***.

6. **Statement of the Key Problem.** The main aim of the present chapter is to elucidate the following questions:

(1) What are the periodic functions with period $2l$ which can be expanded in trigonometric series (11.11), i.e. can be represented as a sum of this kind?

* We remind the reader that the constant A_0 can be regarded as a periodic function with an arbitrary period, in particular, with period T.

(2) How can we determine the coefficients a_0, a_k and b_k, $k = 1, 2, \ldots$, of expansion (11.11) if it is valid?

(3) What is the connection between the character of convergence of series (11.11) and the properties of the function $f(x)$?

7. Orthogonality of Trigonometric System. Fourier Coefficients and Fourier Series. Relation (11.11) is the expansion of the function $f(x)$ into a series with respect to the system of functions

$$\frac{1}{2},\ \cos\frac{\pi x}{l},\ \sin\frac{\pi x}{l},\ \ldots,\ \cos\frac{k\pi x}{l},\ \sin\frac{k\pi x}{l},\ \ldots \tag{11.12}$$

which will be referred to as the **trigonometric system**.

Trigonometric system (11.11) possesses the property that the integral of the product of any two different functions of this system taken over the interval $[-l, l]$ is equal to zero. This property is referred to as the **orthogonality** of the system (the general definition of orthogonal functions will be given in Sec. 1 of § 3). Besides, the integral of the square of every function belonging to (11.11) is unequal to zero.

Indeed, we have

$$\left.\begin{aligned}&\int_{-l}^{l}\frac{1}{2}\cos\frac{k\pi x}{l}\,dx=\frac{1}{2}\,\frac{l}{k\pi}\sin\frac{k\pi x}{l}\Big|_{x=-l}^{x=l}=0\\&\int_{-l}^{l}\frac{1}{2}\sin\frac{k\pi x}{l}\,dx=-\frac{1}{2}\,\frac{l}{k\pi}\cos\frac{k\pi x}{l}\Big|_{x=-l}^{x=l}=\\&\qquad=-\frac{1}{2}\,\frac{l}{k\pi}\left[(-1)^k-(-1)^k\right]=0\end{aligned}\right\} \tag{11.13$_1$}$$

and

$$\int_{-l}^{l}\cos\frac{k\pi x}{l}\sin\frac{n\pi x}{l}\,dx=$$

$$=\frac{1}{2}\int_{-l}^{l}\left[\cos\frac{(k+n)\pi}{l}x+\cos\frac{(k-n)\pi}{l}x\right]dx=0\quad\text{for } k\neq n \tag{11.13$_2$}$$

Similarly,

$$\left.\begin{aligned}&\int_{-l}^{l}\sin\frac{k\pi x}{l}\sin\frac{n\pi x}{l}\,dx=0\quad\text{for } k\neq n\\&\int_{-l}^{l}\sin\frac{k\pi x}{l}\cos\frac{n\pi x}{l}\,dx=0\quad\text{for any } k \text{ and } n\end{aligned}\right\} \tag{11.13$_3$}$$

Finally,

$$\left.\begin{aligned} \int_{-l}^{l} \cos^2 \frac{k\pi x}{l}\,dx &= \int_{-l}^{l} \frac{1+\cos 2\,\frac{k\pi x}{l}}{2}\,dx = l \\ \int_{-l}^{l} \sin^2 \frac{k\pi x}{l}\,dx &= \int_{-l}^{l} \frac{1-\cos 2\,\frac{k\pi x}{l}}{2}\,dx = l \\ \int_{-l}^{l} \left(\frac{1}{2}\right)^2 dx &= \frac{l}{2} \end{aligned}\right\} \qquad (11.13_4)$$

Now let us discuss the problem of determining the coefficients a_0, a_k, b_k, $k=1, 2, \ldots$, of expansion (11.11).

As was proved, a functional series convergent in the mean or uniformly convergent can be integrated term-by-term (see §§ 2 and 6 of Chapter 8). Let us suppose that series (11.11) converges uniformly or in the mean to the function $f(x)$ on the interval $[-l, l]$. It remains convergent (in the same sense) if we multiply both sides of relation (11.11) by any continuous function. This property and the orthogonality of system (11.12) enable us to perform the operations given below and to determine the sought-for coefficients a_0, a_k, b_k, $k=1, 2, \ldots$.

Integrating equality (11.11) termwise we find

$$\int_{-l}^{l} f(x)\,dx = \frac{a_0}{2}\int_{-l}^{l} dx + \sum_{k=0}^{+\infty}\left[a_k \int_{-l}^{l} \cos\frac{k\pi x}{l}\,dx + b_k \int_{-l}^{l} \sin\frac{k\pi x}{l}\,dx\right] = a_0 l$$

whence

$$a_0 = \frac{1}{l}\int_{-l}^{l} f(x)\,dx \qquad (11.14_0)$$

To determine the coefficient a_n in $\cos\frac{n\pi x}{l}$ we multiply equality (11.11) by $\cos\frac{n\pi x}{l}$ and integrate it with respect to x from $-l$ to l. This results (by (11.13$_1$)-(11.13$_4$)) in

$$\int_{-l}^{l} f(x)\cos\frac{n\pi x}{l}\,dx = \frac{a_0}{2}\int_{-l}^{l} \cos\frac{n\pi x}{l}\,dx + \sum_{k=1}^{+\infty} a_k \int_{-l}^{l} \cos\frac{k\pi x}{l}\cos\frac{n\pi x}{l}\,dx +$$
$$+\sum_{k=1}^{+\infty} b_k \int_{-l}^{l} \sin\frac{k\pi x}{l}\cos\frac{n\pi x}{l}\,dx = a_n \int_{-l}^{l} \cos^2\frac{n\pi x}{l}\,dx = a_n l$$

and thus

$$a_n = \frac{1}{l}\int_{-l}^{l} f(x)\cos\frac{n\pi x}{l}\,dx \qquad (11.14_1)$$

Similarly, to find the coefficient b_n in $\sin\frac{n\pi x}{l}$ we multiply equality (11.11) by $\sin\frac{n\pi x}{l}$ and integrate it with respect to x from $-l$ to l, which yields (by (11.13_1)-(11.13_4)) the relation

$$\int_{-l}^{l} f(x)\sin\frac{n\pi x}{l}\,dx = b_n l$$

and thus

$$b_n = \frac{1}{l}\int_{-l}^{l} f(x)\sin\frac{n\pi x}{l}\,dx \qquad (11.14_2)$$

Definition 1. *The numbers a_0, a_n and b_n, $n = 1, 2, \ldots$ determined by formulas (11.14_0), (11.14_1) and (11.14_2) are called the* ***Fourier coefficients of the function*** *$f(x)$* ***with respect to trigonometric system*** (11.12).

Definition 2. *The trigonometric series*

$$\frac{a_0}{2} + \sum_{k=1}^{+\infty}\left(a_k\cos\frac{k\pi x}{l} + b_k\sin\frac{k\pi x}{l}\right) \qquad (11.15)$$

whose coefficients are specified by the function $f(x)$ according to formulas (11.14_0), (11.14_1) and (11.14_2) is called the ***Fourier series of the function*** *$f(x)$.*

It should be noted that the Fourier coefficients a_0, a_k, b_k, $k = 1, 2, \ldots$ determined by formulas (11.14_0), (11.14_1) and (11.14_2) can be defined without imposing the requirement that series (11.11) is convergent. Indeed, for integrals (11.14_0)-(11.14_2) to exist it is sufficient that the function $f(x)$ be integrable on the interval $[-l, l]$. Therefore, to every function $f(x)$ integrable on the interval $[-l, l]$ there corresponds its Fourier series. This correspondence can be written in the form

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{+\infty}\left(a_k\cos\frac{k\pi x}{l} + b_k\sin\frac{k\pi x}{l}\right) \qquad (11.16)$$

where the series on the right-hand side is the trigonometric series whose coefficients are determined by formulas (11.14_0), (11.14_1) and (11.14_2). But in the general case when the function $f(x)$ satisfies only the condition that it is integrable on the interval $[-l, l]$ the

sign of correspondence $\sim$ in relation (11.16) cannot be replaced by the sign of equality. In § 2 we shall introduce some sufficient conditions for expansion (11.11) to be valid.

8. Expanding Even and Odd Functions in Fourier Series. A function $f(x)$ defined on an interval $[-l, l]$ is called even if

$$f(-x) \equiv f(x) \quad \text{for all} \quad x \in [-l, l] \tag{11.17}$$

If a function $f(x)$ defined on an interval $[-l, l]$ satisfies the condition

$$f(-x) \equiv -f(x) \quad \text{for all} \quad x \in [-l, l] \tag{11.18}$$

we say that it is an odd function.

These definitions imply that the graph of an even function is symmetric with respect to the axis of ordinates and the graph of an odd function is symmetric with respect to the origin.

If $f(x)$ is an arbitrary function defined on an interval $[-l, l]$ we can form the functions

$$f_1(x) = \frac{f(x)+f(-x)}{2} \quad \text{and} \quad f_2(x) = \frac{f(x)-f(-x)}{2} \tag{11.19}$$

where the former is even and the latter is odd. We obviously have

$$f(x) = f_1(x) + f_2(x) \quad \text{for all} \quad x \in [-l, l] \tag{11.20}$$

and, consequently, every function $f(x)$ defined on an interval of the form $[-l, l]$ can be represented as the sum of the corresponding even and odd functions.

If $f(x)$ is integrable for $x \in [-l, l]$ we obtain

$$\int_{-l}^{l} f(x)\,dx = \int_{-l}^{0} f(x)\,dx + \int_{0}^{l} f(x)\,dx = \int_{0}^{l} [f(x) + f(-x)]\,dx \tag{11.21}$$

because substituting $-x$ for x we get

$$\int_{-l}^{0} f(x)\,dx = -\int_{l}^{0} f(-x)\,dx = \int_{0}^{l} f(-x)\,dx$$

It follows from relation (11.21) that

$$\int_{-l}^{l} f(x)\,dx = \left\{ \begin{array}{ll} 2\int_{0}^{l} f(x)\,dx & \text{if the function } f(x) \text{ is even} \\ 0 & \text{if the function } f(x) \text{ is odd} \end{array} \right\} \tag{11.22}$$

The functions $\frac{1}{2}$, $\cos\frac{\pi x}{l}$, $\cos\frac{2\pi x}{l}$, $\ldots$, $\cos\frac{k\pi x}{l}$, $\ldots$ are even and the functions $\sin\frac{\pi x}{l}$, $\sin\frac{2\pi x}{l}$, $\ldots$, $\sin\frac{k\pi x}{l}$, $\ldots$ are odd.

Let $f(x)$ be integrable on the interval $[-l, l]$. Then if this function is even its Fourier series is of the form

$$\frac{a_0}{2}+\sum_{k=1}^{+\infty} a_k \cos\frac{k\pi x}{l} \tag{11.23}$$

and if the function $f(x)$ is odd its Fourier series has the form

$$\sum_{k=1}^{\infty} b_k \sin\frac{k\pi x}{l} \tag{11.24}$$

Indeed, if the function $f(x)$ is even the expression $f(x)\cos\frac{k\pi x}{l}$ is also an even function whereas the function $f(x)\sin\frac{k\pi x}{l}$ is odd, and therefore in this case

$$\left.\begin{aligned}
a_0 &= \frac{1}{l}\int_{-l}^{l} f(\xi)\,d\xi = \frac{2}{l}\int_{0}^{l} f(\xi)\,d\xi\\
b_k &= \frac{1}{l}\int_{-l}^{l} f(\xi)\sin\frac{k\pi\xi}{l}\,d\xi = 0, \quad k=1,\ 2,\ \dots\\
a_k &= \frac{1}{l}\int_{-l}^{l} f(\xi)\cos\frac{k\pi\xi}{l}\,d\xi = \frac{2}{l}\int_{0}^{l} f(\xi)\cos\frac{k\pi\xi}{l}\,d\xi,\ k=1,\ 2,\ \dots
\end{aligned}\right\} \tag{11.25}$$

Accordingly, if the function $f(x)$ is odd the function $f(x)\cos\frac{k\pi x}{l}$ is also odd and the function $f(x)\sin\frac{k\pi x}{l}$ is even, and hence

$$\left.\begin{aligned}
&a_0 = \frac{1}{l}\int_{-l}^{l} f(\xi)\,d\xi = 0, \quad a_k = \frac{1}{l}\int_{-l}^{l} f(\xi)\cos\frac{k\pi\xi}{l}\,d\xi = 0\\
&b_k = \frac{1}{l}\int_{-l}^{l} f(\xi)\sin\frac{k\pi\xi}{l}\,d\xi = \frac{2}{l}\int_{0}^{l} f(\xi)\sin\frac{k\pi\xi}{l}\,d\xi,\ k=1,\ 2,\ \dots
\end{aligned}\right\} \tag{11.26}$$

9. Expanding Functions in Fourier Series on the Interval $[-\pi, \pi]$. If it is required to expand a function $f(x)$ defined on the interval $[-\pi, \pi]$ into Fourier's series we put $l=\pi$ in formulas (11.13) and (11.17) and thus obtain the following expressions for the Fourier coefficients and Fourier series:

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\xi)\,d\xi, \quad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\xi)\cos n\xi\,d\xi$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\xi)\sin n\xi\,d\xi, \quad f(x) \sim \frac{a_0}{2}+\sum_{k=1}^{+\infty}(a_k\cos kx + b_k\sin kx)$$

The general case when a function $f(x)$ is originally defined on an interval $[-l, l]$ is reduced to the above by means of the change of the independent variable $x' = \frac{\pi x}{l}$. In fact, if the function $f(x)$ is defined on the interval $[-l, l]$ the new function $\varphi(x') = f\left(\frac{lx}{\pi}\right)$ is defined for $-\pi \leqslant x' \leqslant \pi$. But in view of the applications of Fourier series to the problems of mathematical physics we shall deal with the general case of an arbitrary interval $[-l, l]$ to unify the notation and facilitate the reader to learn the corresponding techniques.

§ 2. FUNDAMENTAL THEOREM ON CONVERGENCE OF FOURIER SERIES

In § 2 our main aim is to prove that Fourier series (11.16) of a periodic piecewise smooth function $f(x)$ with period $2l$ converges to $f(x)$ at each point of continuity of $f(x)$.

1. Class of Piecewise Smooth Functions. *A function $f(x)$ is said to be* ***piecewise continuous on an interval*** *$[a, b]$ if it is continuous everywhere on this interval except possibly at a finite number of points of ordinary discontinuity (i.e. discontinuity of the first kind or jump discontinuity).*

Such a function has finite left-hand and right-hand limits

$$f(x-0) = \lim_{\substack{z\to 0\\ z<0}} f(x+z), \quad f(x+0) = \lim_{\substack{z\to 0\\ z>0}} f(x+z) \tag{11.27}$$

at each interior point x of the interval $[a, b]$ and finite limits $f(a+0)$ and $f(b-0)$ at the end points a and b of the interval $[a, b]$. At each point of continuity of $f(x)$ we have $f(x-0) = f(x+0) = f(x)$, and thus $f(x-0)$ and $f(x+0)$ do not coincide only at a finite number of points at which $f(x)$ has the jumps $f(x+0) - f(x-0) \neq 0$.

A piecewise continuous function $f(x)$ defined on an interval $[a, b]$, $a < b$, is said to be ***piecewise smooth*** *if the derivative $f'(x)$ exists and is continuous everywhere on $[a, b]$ except possibly at a finite number of points at which finite right-hand and left-hand limits*

$$f'(x+0) = \lim_{\substack{z\to 0\\ z>0}} f'(x+z), \quad f'(x-0) = \lim_{\substack{z\to 0\\ z>0}} f'(x-z) \tag{11.28}$$

exist. It is also supposed that finite limits $f'(a+0)$ and $f'(b-0)$ exist at the end points of the interval $[a, b]$.

A piecewise smooth function $f(x)$ possesses finite **left-hand** and **right-hand derivatives**

$$\left.\begin{aligned} f'_L(x) &= \lim_{\substack{z\to 0\\ z>0}} \frac{f(x-z)-f(x-0)}{-z} \\ f'_R(x) &= \lim_{\substack{z\to 0\\ z>0}} \frac{f(x+z)-f(x+0)}{z} \end{aligned}\right\} \qquad (11.29)$$

at each point x of the interval $[a, b]$. In fact, applying the formula of finite increments* and using relation (11.28) we obtain

$$\lim_{\substack{z\to 0\\ z>0}} \frac{f(x\pm z)-f(x\pm 0)}{\pm z} = \lim_{\substack{z\to 0\\ z>0}} f'(x\pm\theta z) = f'(x\pm 0) \qquad (11.30)$$

and hence the derivatives $f'_L(x)$ and $f'_R(x)$ exist and besides the relations

$$f'_L(x) = f'(x-0) \quad \text{and} \quad f'_R(x) = f'(x+0)$$

hold.

The graph of a piecewise smooth function $f(x)$ possesses a uniquely specified tangent line at each point except possibly at a finite

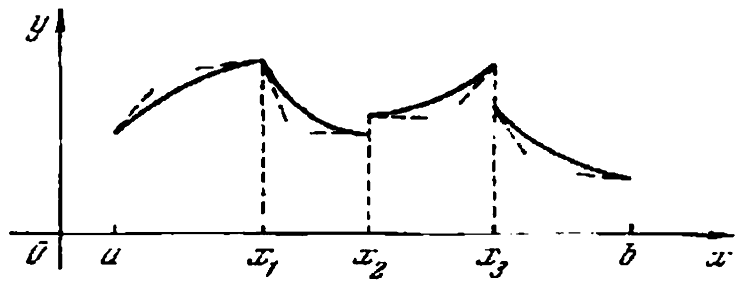


Fig. 11.2

number of points but at these there must exist uniquely specified left-hand and right-hand tangent lines (see Fig. 11.2).

If $f(x)$ is a piecewise smooth function on $[a, b]$ we can break up the interval $[a, b]$ into a finite number of subintervals

$$[a_0, a_1], \ [a_1, a_2], \ \ldots, \ [a_i, a_{i+1}], \ \ldots, \ [a_N, a_{N+1}]$$

where

$$a_0 = a < a_1 < \ldots < a_i < a_{i+1} < \ldots < a_N < a_{N+1} = b$$

---

* Here we apply the formula of finite increments written in the form $f(x+z) - f(x+0) = zf'(x+\theta z)$ where $0<\theta<1$ which can be derived in the following manner: we take δ, $0<\delta<z$, and write down the ordinary formula of finite increments

$$f(x+z) - f(x+\delta) = (z-\delta)\, f'(x+\xi)$$

where $\delta < \xi < z$ and then pass to the limit as $\delta \to 0$.

such that the functions $f(x)$ and $f'(x)$ are continuous at the interior points of each interval $[a_i, a_{i+1}]$, $i = 0, \ldots, N$, and tend to finite limits

$$f(a_i + 0),\ f'(a_i + 0) \quad \text{and} \quad f(a_{i+1} - 0),\ f'(a_{i+1} - 0)$$

when x approaches a_i from the right and a_{i+1} from the left. It follows that the functions $f(x)$ and $f'(x)$ are bounded on each interval $[a_i, a_{i+1}]$ and, consequently, they are bounded on the whole interval $[a, b]$.*

2. Formulation of Fundamental Theorem on Convergence of Fourier Series.

Theorem 11.1. If $f(x)$ is a piecewise smooth function on an interval $-l \leqslant x \leqslant l$, its Fourier series converges at each point x of the interval and the sum

$$S(x) = \frac{a_0}{2} + \sum_{k=1}^{+\infty} \left(a_k \cos \frac{k\pi x}{l} + b_k \sin \frac{k\pi x}{l}\right) \qquad (11.31)$$

of this series satisfies the following relations:

(1) $S(x) = f(x)$ *if* $-l < x < l$ *and* $f(x)$ *is continuous at the point* x,

(2) $S(x) = \dfrac{f(x+0) + f(x-0)}{2}$ *if* $-l < x < l$ *and* x *is a point of discontinuity of* $f(x)$,

(3) $S(-l) = S(l) = \dfrac{f(-l+0) + f(l-0)}{2}$.

Note. If $f(x)$ is continuous at a point x, $-l < x < l$, we have $f(x-0) = f(x+0) = f(x)$ and, consequently,

$$\frac{f(x+0) + f(x-0)}{2} = \frac{2f(x)}{2} = f(x)$$

Therefore equalities (1) and (2) can be replaced by the relation

$$S(x) = \frac{f(x+0) + f(x-0)}{2} \qquad (11.32)$$

which holds for every interior point x of the interval $[-l, l]$.

3. Key Lemma. To prove the theorem we need the following lemma

---

* In fact, if we redefine $f(x)$ and $f'(x)$ at the end points of an interval $[a_i, a_{i+1}]$, $i = 0, 1, \ldots, N$, by putting $f(a_i) = f(a_i + 0)$, $f(a_{i+1}) = f(a_{i+1} - 0)$, $f'(a_i) = f'(a_i + 0)$ and $f'(a_{i+1}) = f'(a_{i+1} - 0)$, the functions $f(x)$ and $f'(x)$ become continuous on the closed interval $[a_i, a_{i+1}]$ and, consequently, bounded on it. But, in this construction, the values of $f(x)$ and $f'(x)$ may only change at the end points of the interval $[a_i, a_{i+1}]$ and hence the original functions $f(x)$ and $f'(x)$ are also bounded on $[a_i, a_{i+1}]$.

Key Lemma. If $f(x)$ is a piecewise smooth function on an interval $a \leqslant x \leqslant b$ we have

$$\lim_{\alpha \to \infty} \int_a^b f(x) \sin \alpha x \, dx = 0 \tag{11.33}$$

Proof. Let us divide $[a, b]$ into parts

$$[a_0, a_1], [a_1, a_2], \ldots, [a_i, a_{i+1}], \ldots, [a_N, a_{N+1}]$$

where $a = a_0 < a_1 < \ldots < a_i < a_{i+1} < \ldots < a_N < a_{N+1} = b$ in such a way that the functions $f(x)$ and $f'(x)$ are continuous in the interior of each subinterval $[a_i, a_{i+1}]$, $i = 0, 1, \ldots, N$, and tend to finite limits

$$f(a_i + 0), \ f'(a_i + 0) \quad \text{and} \quad f(a_{i+1} - 0), \ f'(a_{i+1} - 0)$$

when x tends to a_{i+1} from the left and to a_i from the right, respectively. Since

$$\int_a^b f(x) \sin \alpha x \, dx = \sum_{i=0}^{N} \int_{a_i}^{a_{i+1}} f(x) \sin \alpha x \, dx \tag{11.34}$$

it is sufficient to prove that

$$\lim_{\alpha \to \infty} \int_{a_i}^{a_{i+1}} f(x) \sin \alpha x \, dx = 0 \tag{11.35}$$

for $0 \leqslant i \leqslant N$. We can regard $f(x)$ and $f'(x)$ as being continuous on the closed interval $[a_i, a_{i+1}]$ (see footnote on page 488), and therefore it is allowable to integrate by parts, which gives

$$\int_{a_i}^{a_{i+1}} f(x) \sin \alpha x \, dx = -\frac{f(x) \cos \alpha x}{\alpha} \Bigg|_{x=a_i+0}^{x=a_{i+1}-0} + \frac{1}{\alpha} \int_{a_i}^{a_{i+1}} f'(x) \cos \alpha x \, dx \tag{11.36}$$

The functions $f(x)$ and $f'(x)$ being bounded on $[a, b]$, there exist constants M and M' such that $|f(x)| \leqslant M$ and $|f'(x)| \leqslant M'$ everywhere on $[a, b]$. Hence, relation (11.36) implies the inequality

$$\left| \int_{a_i}^{a_{i+1}} f(x) \sin \alpha x \, dx \right| \leqslant \frac{2M}{\alpha} + \frac{M'(a_{i+1} - a_i)}{\alpha} \tag{11.37}$$

Passing to the limit in equality (11.37) for $\alpha \to \infty$ we obtain relation (11.35). The key lemma has thus been proved.

Note 1. It can be shown that the above lemma applies to a considerably wider class of functions. For instance, if $f(x)$ is an absolutely integrable function on $[a, b]$, that is the integral $\int_a^b |f(x)|\,dx < +\infty$ (which may be improper) exists, the lemma remains valid.

Note 2. We can similarly take $\cos \alpha x$ under the sign of integration instead of $\sin \alpha x$.

4. Proof of Convergence Theorem. Let $f(x)$ be a piecewise continuous and piecewise smooth function on an interval $-l \leqslant x \leqslant l$. We shall extend the function $f(x)$ as a periodic function with period $2l$ from the interval $[-l, l]$ to the entire x-axis and prove that

$$\left\{S_n(x) - \frac{f(x-0)+f(x+0)}{2}\right\} \to 0 \quad \text{for } n \to +\infty \tag{11.38}$$

for every x where

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos\frac{k\pi x}{l} + b_k \sin\frac{k\pi x}{l}\right) \tag{11.39}$$

is the nth partial sum of the Fourier series of the function $f(x)$ corresponding to the interval $-l \leqslant x \leqslant l$.* Substituting the expressions of the Fourier coefficients

$$\left.\begin{aligned} a_0 &= \frac{1}{l}\int_{-l}^{l} f(\xi)\,d\xi \\ a_k &= \frac{1}{l}\int_{-l}^{l} f(\xi)\cos\frac{k\pi\xi}{l}\,d\xi, \quad k = 1, 2, \ldots \\ b_k &= \frac{1}{l}\int_{-l}^{l} f(\xi)\sin\frac{k\pi\xi}{l}\,d\xi, \quad k = 1, 2, \ldots \end{aligned}\right\} \tag{11.40}$$

* After the periodic extension has been performed, $f(x)$ becomes a periodic function with period $2l$, and the functions $\frac{1}{2}$, $\cos\frac{\pi x}{l}$, $\sin\frac{\pi x}{l}$, ..., $\cos\frac{k\pi x}{l}$, $\sin\frac{k\pi x}{l}$, ... are also periodic with period $2l$. Therefore the integrals appearing when we compute the Fourier coefficients $a_0 = \frac{1}{l}\int_{-l}^{l} f(x)\,dx$, $a_n = \frac{1}{l}\int_{-l}^{l} f(x)\cos\frac{n\pi x}{l}\,dx$, $b_n = \frac{1}{l}\int_{-l}^{l} f(x)\sin\frac{n\pi x}{l}\,dx$, $n = 1, 2, \ldots$, can be taken not only over the original interval $[-l, l]$ but also over any other interval of length $2l$, which does not affect the values of these coefficients.

into (11.39) we obtain

$$S_n(x)=\frac{1}{2l}\int_{-l}^{l} f(\xi)\,d\xi+\frac{1}{l}\sum_{k=1}^{n}\int_{-l}^{l} f(\xi)\left[\cos\frac{k\pi\xi}{l}\cos\frac{k\pi x}{l}+\right.$$

$$\left.+\sin\frac{k\pi\xi}{l}\sin\frac{k\pi x}{l}\right]d\xi=\frac{1}{l}\int_{-l}^{l} f(\xi)\left[\frac{1}{2}+\sum_{k=1}^{n}\cos\frac{k\pi(\xi-x)}{l}\right]d\xi=$$

$$=\frac{1}{l}\int_{-l-x}^{l-x} f(x+z)\left[\frac{1}{2}+\sum_{k=1}^{n}\cos\frac{k\pi z}{l}\right]dz \qquad (11.41)$$

where $z=\xi-x$. Now let us derive an expression for the sum

$$\sigma_n(z)=\frac{1}{2}+\sum_{k=1}^{n}\cos\frac{k\pi z}{l} \qquad (11.42)$$

Miltiplying both sides of (11.42) by $2\sin\frac{\pi z}{2l}$ we obtain

$$2\sigma_n(z)\sin\frac{\pi z}{2l}=\sin\frac{\pi z}{2l}+\sum_{k=1}^{n}2\sin\frac{\pi z}{2l}\cos\frac{k\pi z}{l}=$$

$$=\sin\frac{\pi z}{2l}+\sum_{k=1}^{n}\left[\sin\left(k+\frac{1}{2}\right)\frac{\pi z}{l}-\sin\left(k-\frac{1}{2}\right)\frac{\pi z}{l}\right]=\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}$$

whence

$$\sigma_n(z)=\frac{1}{2}+\sum_{k=1}^{n}\cos\frac{k\pi z}{l}=\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}} \qquad (11.43)$$

Substituting (11.43) into (11.41) we arrive at the following formula for the nth partial sum of the Fourier series:

$$S_n(x)=\frac{1}{l}\int_{-l-x}^{l-x} f(x+z)\,\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz \qquad (11.44)$$

The function $f(x)$ (after it has been periodically extended from the interval $[-l, l]$ to the whole x-axis) is periodic with period $2l$, and the expression $\dfrac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}$ is, by (11.43), a periodic function of z with period $2l$. Therefore the product of $f(x+z)$

by $\dfrac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}$ is also a periodic function of z with period $2l$, and the integral of this product taken over any interval of the z-axis of length $2l$ has one and the same value. Consequently, the limits of integration $-l-x$ and $l-x$ in integral (11.44) (which is taken over an interval of length $2l$) can be replaced by the limits $-l$ and l. This results in

$$S_n(x)=\frac{1}{l}\int_{-l}^{l} f(x+z)\,\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz \tag{11.45}$$

Integrating (11.43) with respect to z from $-l$ to l we deduce

$$\frac{1}{l}\int_{-l}^{l}\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz=\frac{1}{2l}\int_{-l}^{l}dz=1 \tag{11.46}$$

since $\int_{-l}^{l}\cos\frac{k\pi z}{l}\,dz=0$ for $k=1,\ 2,\ 3,\ \ldots$. But the expression $\dfrac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}$ is an even function of z (see § 1, Sec. 8), and consequently (11.46) implies that

$$\frac{1}{l}\int_{-l}^{0}\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz=\frac{1}{2}\quad\text{and}\quad\frac{1}{l}\int_{0}^{l}\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz=\frac{1}{2} \tag{11.47}$$

Multiplying the first equality (11.47) by $f(x-0)$ and the second by $f(x+0)$ and adding together the results we obtain

$$\frac{f(x-0)+f(x+0)}{2}=\frac{1}{l}\int_{-l}^{0} f(x-0)\,\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz+$$

$$+\frac{1}{l}\int_{0}^{l} f(x+0)\,\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz \tag{11.48}$$

Now subtracting (11.48) from (11.45) we get

$$S_n(x)-\frac{f(x-0)+f(x+0)}{2}=$$

$$=\frac{1}{l}\int_{-l}^{0}[f(x-z)-f(x-0)]\,\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz+$$

$$+\frac{1}{l}\int_{0}^{l}[f(x+z)-f(x+0)]\,\frac{\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}}{2\sin\frac{\pi z}{2l}}\,dz \qquad (11.49)$$

Let us prove that both integrals on the right-hand side of equality (11.49) tend to zero when $n\to+\infty$. For instance, let us take the second integral. We can represent the integral in question in the form

$$J_n=\frac{1}{\pi}\int_{0}^{\delta}\frac{f(x+z)-f(x+0)}{z}\,\frac{\frac{\pi z}{2l}}{\sin\frac{\pi z}{2l}}\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}\,dz+$$

$$+\frac{1}{l}\int_{\delta}^{l}\frac{f(x+z)-f(x+0)}{2\sin\frac{\pi z}{2l}}\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}\,dz=J_n'+J_n'' \quad (11.50)$$

where $0<\delta<l$. Suppose we are given an arbitrary $\varepsilon>0$. Let us show that for sufficiently small $\delta>0$ the first integral in (11.50) is less than $\frac{\varepsilon}{2}$ in its modulus for all $n=1, 2, \ldots$. In fact, $\frac{f(x+z)-f(x+0)}{z}\to f'(x+0)$ as $z\to 0+0$*, and therefore, for a sufficiently small $\delta>0$ and all z belonging to the interval $0<z<\delta$, we have the inequality

$$\left|\frac{f(x+z)-f(x+0)}{z}\right|<|f'(x+0)|+1$$

Furthermore, $\dfrac{\frac{\pi z}{2l}}{\sin\frac{\pi z}{2l}}\to 1$ for $z\to 0$ and, consequently, for a sufficiently small $\delta>0$ and all z from the interval $0<z<\delta$ we have the inequality

$$1<\frac{\frac{\pi z}{2l}}{\sin\frac{\pi z}{2l}}<2$$

---

* See relation (11.30).

Finally, for all real z and all $n=1, 2, \ldots$ we can write

$$\left|\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}\right|\leqslant 1$$

Thus, the relation

$$|J_n'|\leqslant\frac{1}{\pi}\int_0^\delta\left|\frac{f(x+z)-f(x+0)}{z}\right|\left|\frac{\frac{\pi z}{2l}}{\sin\frac{\pi z}{2l}}\right|\left|\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}\right|dz\leqslant$$

$$\leqslant\frac{2\delta}{\pi}[|f'(x+0)|+1]$$

is fulfilled for all $n=1, 2, \ldots$ provided that $\delta>0$ is sufficiently small. Taking a sufficiently small $\delta>0$ such that $\frac{2\delta}{\pi}[|f'(x+0)|+1]<\frac{\varepsilon}{2}$ we see that

$$|J_n'|<\frac{\varepsilon}{2}\ \text{for all}\quad n=1, 2, \ldots \tag{11.51}$$

We now fix a value of $\delta>0$ thus determined and proceed to estimate the second integral in (11.50). We have

$$J_n''=\frac{1}{l}\int_\delta^l\frac{f(x+z)-f(x+0)}{2\sin\frac{\pi z}{2l}}\sin\left(n+\frac{1}{2}\right)\frac{\pi z}{l}\,dz$$

where the function $\dfrac{f(x+z)-f(x+0)}{2\sin\frac{\pi z}{2l}}$ is piecewise continuous and piecewise smooth on the interval $\delta\leqslant z\leqslant l$ (for $\delta>0$) because here the numerator is piecewise smooth and the denominator is a continuously differentiable function which does not vanish on that interval. Then, by the key lemma, we have $J_n''\to 0$ for $n\to+\infty$ and, consequently, for all sufficiently large values of n we obtain the inequality

$$|J_n''|<\frac{\varepsilon}{2} \tag{11.52}$$

Taking into account (11.50) we conclude, on the basis of (11.51) and (11.52), that

$$|J_n|\leqslant|J_n'|+|J_n''|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon \tag{11.53}$$

for all sufficiently large n which means that $\lim\limits_{n\to+\infty}J_n=0$.

We similarly prove that the first integral on the right-hand side of (11.49) also tends to zero for $n\to+\infty$, and thus

$$\lim_{n\to+\infty}S_n(x)=\frac{f(x+0)+f(x-0)}{2} \tag{11.54}$$

We remind the reader that the function $f(x)$ is regarded as being periodically extended, with period $2l$, from the interval $[-l, l]$ to the whole x-axis. Therefore

$$f(l+0) = f(-l+0) \quad \text{and} \quad f(-l-0) = f(l-0) \tag{11.55}$$

Now substituting $x = -l$ and $x = l$ into (11.54) and taking advantage of relation (11.45) we obtain

$$\lim_{n\to+\infty} S_n(-l) = \lim_{n\to+\infty} S_n(l) = \frac{f(-l+0)+f(l-0)}{2} \tag{11.56}$$

and thus the proof of the theorem has been completed.

5. Examples. 1. Let us take the function $f(x) = x$ regarded as being defined on the interval $[-l, l]$ and expand it into a Fourier series. This function is odd, and hence

$$a_0=0, \quad a_n=0, \quad b_n=\frac{2}{l}\int_0^l f(x)\sin\frac{n\pi x}{l}\,dx, \qquad n=1, 2, \ldots$$

Integrating by parts we obtain

$$b_n=\frac{2}{l}\int_0^l x\sin\frac{n\pi x}{l}\,dx=$$

$$=\frac{2}{l}\left\{\left(-\frac{xl}{n\pi}\cos\frac{n\pi x}{l}\right)\Big|_{x=0}^{x=l}+\frac{l}{n\pi}\int_0^l\cos\frac{n\pi x}{l}\,dx\right\}=\frac{2l}{n\pi}(-1)^{n+1}$$

Consequently, according to the convergence theorem, we have

$$S(x)=\frac{2l}{\pi}\sum_{k=1}^{+\infty}\frac{(-1)^{k+1}}{k}\sin\frac{k\pi x}{l}=$$

$$=\begin{cases}\dfrac{f(x-0)+f(x+0)}{2}=x & \text{for} \quad -l<x<l\\[2ex] \dfrac{f(-l+0)+f(l-0)}{2}=\dfrac{-l+l}{2}=0 & \text{for} \quad x=\pm l\end{cases} \tag{11.57}$$

The graphs of $f(x)=x$ and $S(x)$ are shown in Fig. 11.3.

The function $S(x)$ is periodic with period $2l$, and $S(x)=x$ only for $-l<x<l$. At the points $x=(2k+1)l$, $k=0, \pm1, \pm2, \ldots$, the function $S(x)$ has discontinuities, its left-hand and right-hand limits being, respectively, equal to $f[(2k+1)l-0]=-1$ and $f[(2k+1)l+0]=1$, $k=0, \pm1, \pm2, \ldots$, at these points. At every point $x=(2k+1)l$, $k=0, \pm1, \pm2, \ldots$, the function $S(x)$ assumes the value which is the half-sum of the left-hand and right-hand limits, i.e. equal to zero. In Fig. 11.4

we see the graph of the partial sum $S_5(x) = \frac{2l}{\pi} \sum_{k=1}^{5} \frac{(-1)^{k+1}}{k} \sin \frac{k\pi x}{l}$ constructed on the interval $-l \leqslant x \leqslant l$.

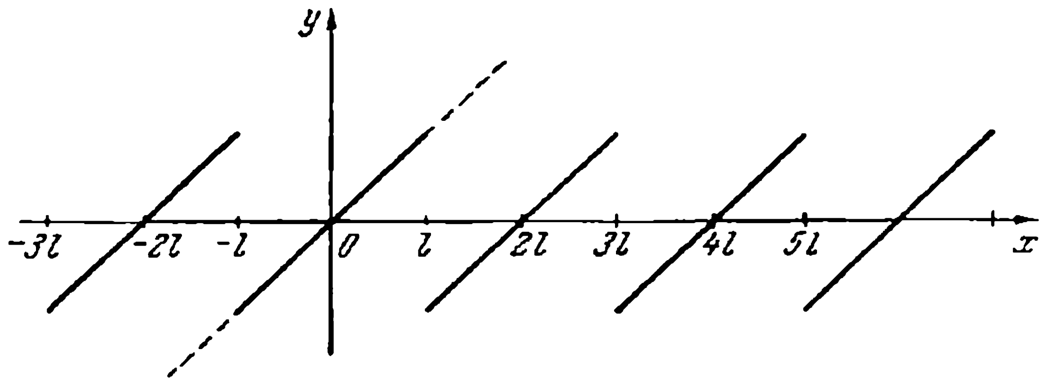


Fig. 11.3

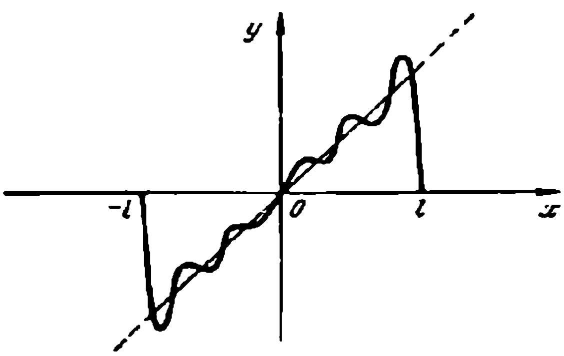


Fig. 11.4

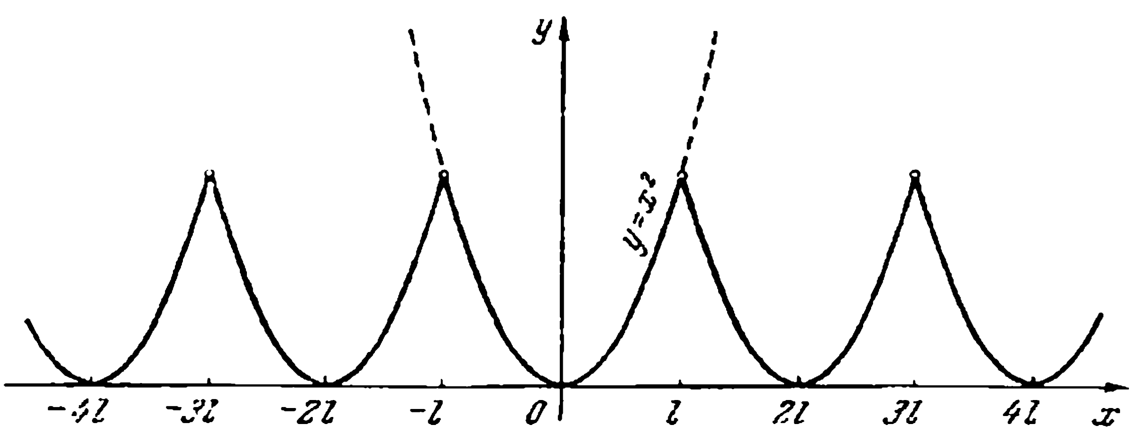


Fig. 11.5

2. Let us take the function $f(x) = x^2$ on the interval $[-l, l]$. This function being even, we have

$$a_0 = \frac{2}{l} \int_0^l x^2\,dx, \quad a_n = \frac{2}{l} \int_0^l x^2 \cos \frac{n\pi x}{l}\,dx, \quad b_n = 0, \qquad n = 1, 2, \ldots$$

After the coefficients a_0 and a_n, $n = 1, 2, \ldots$, have been computed, we find, by the convergence theorem, that

$$S(x) = \frac{l^2}{3} + \frac{4l^2}{\pi^2} \sum_{k=1}^{+\infty} \frac{(-1)^k}{k^2} \cos \frac{k\pi x}{l} = x^2 \quad \text{for} \quad -l \leqslant x \leqslant l \qquad (11.58)$$

The graphs of the function $f(x) = x^2$ and of the sum $S(x)$ of Fourier's series (11.58) are depicted in Fig. 11.5. The function $S(x)$ is continuous everywhere including the points $x = (2k + 1)\pi$, $k = 0, \pm 1, \pm 2, \ldots$ since for the function $f(x) = x^2$ we have $f(-l + 0) = f(l - 0) = l^2$.

If we apply this expansion to the interval $[-\pi, \pi]$, that is for $l = \pi$, equality (11.58) takes the form

$$x^2 = \frac{\pi^2}{3} + 4 \sum_{k=1}^{+\infty} \frac{(-1)^k}{k^2} \cos kx, \qquad -\pi \leqslant x \leqslant \pi \tag{11.59}$$

Substituting $x = 0$ into (11.59) we obtain the useful relation

$$\frac{\pi^2}{12} = 1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \ldots + (-1)^{k-1} \frac{1}{k^2} + \ldots \tag{11.60}$$

3. Let $f(x)$ be defined on the interval $[-l, l]$ by the relations

$$f(x) \begin{cases} c_1 & \text{for} \quad -l < x < 0 \\ c_2 & \text{for} \quad 0 < x < l \end{cases}$$

Then

$$a_0 = \frac{1}{l} \int_{-l}^{l} f(x)\, dx = \frac{1}{l} \int_{-l}^{0} c_1\, dx + \frac{1}{l} \int_{0}^{l} c_2\, dx = c_1 + c_2$$

$$a_n = \frac{1}{l} \int_{-l}^{l} f(x) \cos \frac{n\pi x}{l}\, dx = \frac{1}{l} \int_{-l}^{0} c_1 \cos \frac{n\pi x}{l}\, dx +$$

$$+ \frac{1}{l} \int_{0}^{l} c_2 \cos \frac{n\pi x}{l}\, dx = 0, \qquad n = 1, 2, \ldots$$

and

$$b_n = \frac{1}{l} \int_{-l}^{l} f(x) \sin \frac{n\pi x}{l}\, dx = \frac{1}{l} \int_{-l}^{0} c_1 \sin \frac{n\pi x}{l}\, dx +$$

$$+ \frac{1}{l} \int_{0}^{l} c_2 \sin \frac{n\pi x}{l}\, dx, \qquad n = 1, 2, \ldots$$

Consequently, $b_n = 0$ for even n and $b_n = \frac{2(c_2 - c_1)}{\pi n}$ for odd n. There-

fore, by the convergence theorem, we obtain

$$S(x)=\frac{c_1+c_2}{2}+\frac{2(c_2-c_1)}{\pi}\left\{\sin\frac{\pi x}{l}+\frac{1}{3}\sin\frac{3\pi x}{l}+\frac{1}{5}\sin\frac{5\pi x}{l}+\ldots\right\}=$$

$$=\begin{cases} c_1 & \text{for} \quad -l<x<0 \\ c_2 & \text{for} \quad 0<x<l \\ \frac{c_1+c_2}{2} & \text{for} \quad x=\pm l,\ x=0 \end{cases}$$

The graph of $S(x)$ is shown in Fig. 11.6.

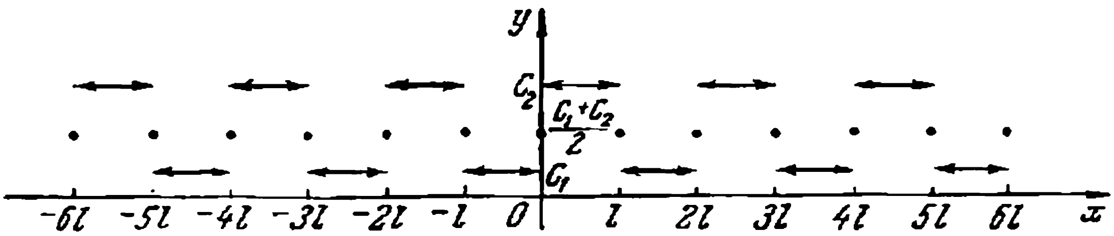


Fig. 11.6

In Fig. 11.7 we see the graphs of the partial sums $S_1(x)$, $S_2(x)$ and $S_3(x)$ for the case $c_1=-1$ and $c_2=+1$. In this case the series turns into

$$S(x)=\frac{4}{\pi}\left\{\sin\frac{\pi x}{l}+\frac{1}{3}\sin\frac{3\pi x}{l}+\frac{1}{5}\sin\frac{5\pi x}{l}+\ldots\right\}=$$

$$=\begin{cases} -1 & \text{for} \quad -l<x<0 \\ 1 & \text{for} \quad 0<x<l \\ 0 & \text{for} \quad x=0,\ x=\pm l \end{cases}$$

6. Fourier Sine and Cosine Series for Functions Defined on Interval $[0, l]$. Let a piecewise continuous and piecewise smooth function $f(x)$ be defined on an interval $0\leqslant x\leqslant l$. It can be extended in various ways to the interval $-l\leqslant x\leqslant 0$, in particular, (1) as an even function or (2) as an odd function. In the former case we obtain an even function on the interval $[-l, l]$, for which

$$a_0=\frac{2}{l}\int\limits_0^l f(\xi)\,d\xi,\quad a_k=\frac{2}{l}\int\limits_0^l f(\xi)\cos\frac{k\pi\xi}{l}\,d\xi,\quad b_k=0,\qquad k=1,\ 2,\ \ldots \tag{11.61}$$

and whose Fourier series, on the interval $[-l, l]$, is of the form

$$f(x)\sim\frac{a_0}{2}+\sum_{k=1}^{+\infty}a_k\cos\frac{k\pi x}{l} \tag{11.62}$$

In the latter case we have an odd function on the interval $[-l, l]$ with Fourier's coefficients

$$a_0 = 0, \quad a_k = 0, \quad b_k = \frac{2}{l}\int_0^l f(\xi) \sin \frac{k\pi\xi}{l}\, d\xi \tag{11.63}$$

and Fourier's series

$$f(x) \sim \sum_{k=1}^{+\infty} b_k \sin \frac{k\pi x}{l} \tag{11.64}$$

Each of the series (11.62) and (11.64) converges to $f(x)$ on the interval $0 < x < l$ at the points of continuity of the function $f(x)$.

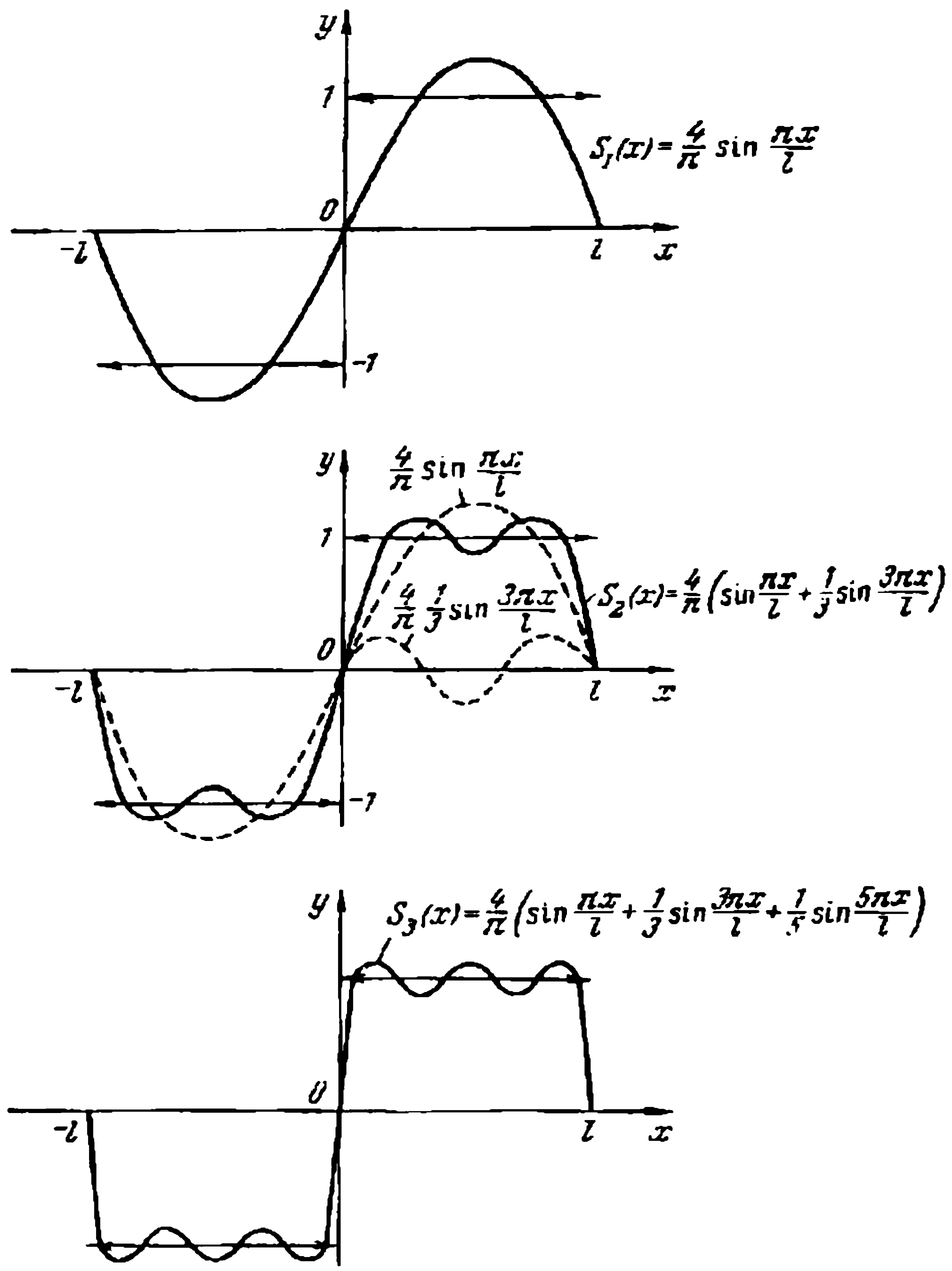


Fig. 11.7

Thus, we see that every piecewise smooth function $f(x)$ defined on an interval $0 \leqslant x \leqslant l$ can be expanded both into a series of form (11.62) involving only cosines whose coefficients are determined

by formulas (11.61) and into a series of type (11.64) containing only sines with coefficients found by formulas (11.63). Series (11.62) and (11.64) are termed, respectively, a **(Fourier) cosine series** and **sine series** (they are also referred to as **Fourier's half-range or incomplete series**).

For example, let $f(x) = x$ on the interval $0 \leqslant x \leqslant l$. If $f(x)$ is extended "in odd fashion" we arrive at the function $f(x) = x$ defined on the interval $-l \leqslant x \leqslant l$ whose Fourier series has already

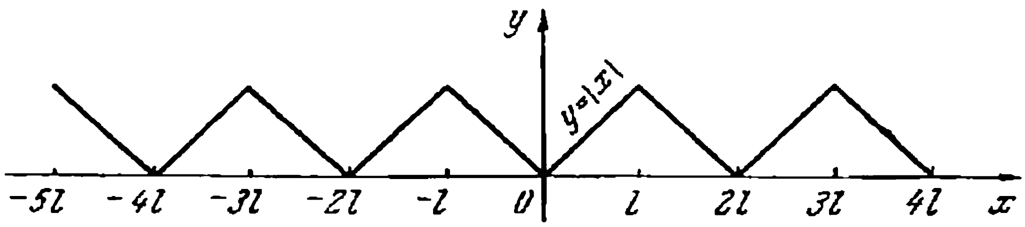


Fig. 11.8

been studied (see Example 1 and Fig. 11.3). Extending $f(x)$ "in even fashion" we obtain the function $f(x) = |x|$ on the interval $-l \leqslant x \leqslant l$. Expanding $f(x) = |x|$ into a cosine series on the interval $0 \leqslant x \leqslant l$ we obtain, by formulas (11.61) and (11.62), the relation

$$x = \frac{a_0}{2} + \sum_{k=1}^{+\infty} a_k \cos\frac{k\pi x}{l} \quad \text{for} \quad 0 \leqslant x \leqslant l$$

where

$$a_0 = \frac{2}{l}\int_0^l x\,dx = l, \quad a_n = \frac{2}{l}\int_0^l x\cos\frac{k\pi x}{l}\,dx =$$

$$= \frac{2l}{\pi^2 n^2}[(-1)^n - 1] = \begin{cases} 0 & \text{for even } n \\ -\dfrac{4l}{\pi^2 n^2} & \text{for odd } n \end{cases}$$

Consequently,

$$x = \frac{l}{2} - \frac{4l}{\pi^2}\left\{\cos\frac{\pi x}{l} + \frac{1}{3^2}\cos\frac{3\pi x}{l} + \frac{1}{5^2}\cos\frac{5\pi x}{l} + \dots\right\}$$
$$\text{for} \quad 0 \leqslant x \leqslant l \qquad (11.65)$$

The validity of equality (11.65) for $x = 0$ and $x = l$ can easily be established if we regard (11.65) as Fourier's series of the even function $f(x) = |x|$ defined on the interval $[-l, l]$. Such an interpretation of series (11.65) enables us to construct the graph of its sum for any values of x (see Fig. 11.8).

Generally, when an arbitrary piecewise continuous and piecewise smooth function $f(x)$ defined on an interval $[0, l]$ is extended in even fashion to the interval $[-l, 0]$ we always have

$$f(-0) = f(+0) \quad \text{and} \quad f(-l+0) = f(l-0) \qquad (11.66)$$

Consequently, the sum of the corresponding Fourier series is continuous at the points $x = 0$ and $x = \pm l$, and its values at these points are, respectively, equal to $f(+0) = f(-0)$ and $f(-l+0) = f(l-0)$. If, in addition, the original function $f(x)$ is continuous at the end points of the interval $[0, l]$, that is $f(+0) = f(0)$ and $f(l-0) = f(l)$, the sum of its cosine series is equal to $f(x)$ at these end points.

On the contrary, if we extend a function $f(x)$ (originally defined for $0 \leqslant x \leqslant l$) in odd fashion to the interval $-l \leqslant x \leqslant 0$, the sum of the corresponding sine series may have discontinuities at the

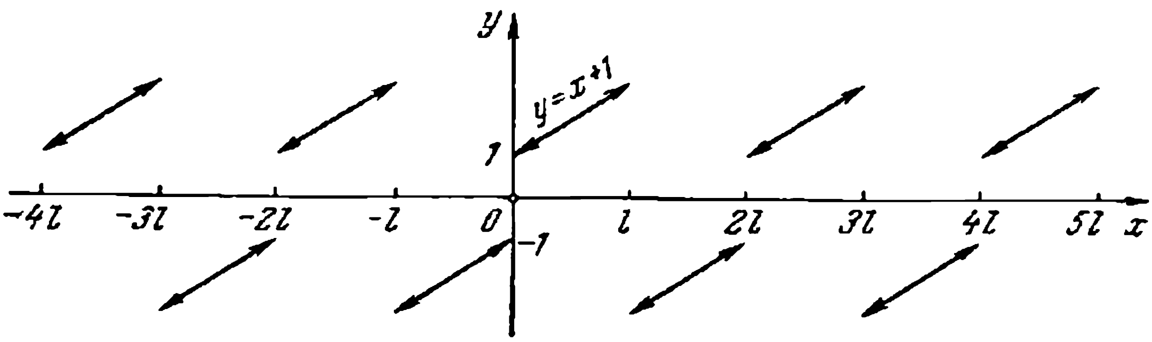


Fig. 11.9

points $x = 0$ and $x = \pm l$ despite the continuity and smoothness of $f(x)$ on the interval $0 \leqslant x \leqslant l$. In Fig. 11.9 we see the graph of the sum of the Fourier series corresponding to the function $f(x) = x + 1$ originally defined for $0 \leqslant x \leqslant l$ and then extended in odd fashion. In the general case, when a function $f(x)$ is extended as an odd function from an interval $[0, l]$ to the interval $[-l, 0]$ we always have $f(-0) = -f(+0)$ and $f(-l+0) = -f(l-0)$. But for the sum of the corresponding Fourier series to be continuous at the points $x = 0$ and $x = \pm l$ it is necessary that the relations $f(-0) = f(+0)$ and $f(-l+0) = f(l-0)$ hold, and hence this sum is continuous only if

$$f(+0) = 0 \quad \text{and} \quad f(l-0) = 0 \tag{11.67}$$

§ 3. FOURIER SERIES WITH RESPECT TO GENERAL ORTHOGONAL SYSTEMS. BESSEL'S INEQUALITY

Fourier's trigonometric series expansion of a function $f(x)$ is a special case of expansion of $f(x)$ in a series with respect to an *orthogonal system of functions*.

1. Orthogonal Systems of Functions. *Two functions $\varphi(x)$ and $\psi(x)$ integrable* on an interval $[a, b]$ are said to be orthogonal on*

* In what follows, unless otherwise stated, integrability is understood in the sense of proper integral and the functions are supposed to be real

$[a, b]$ if

$$\int_a^b \varphi(x)\,\psi(x)\,dx = 0 \tag{11.68}$$

A system of functions

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots \tag{11.69}$$

integrable on $[a, b]$ is called ***orthogonal*** *on $[a, b]$ if*

$$\int_a^b \varphi_i(x)\,\varphi_k(x)\,dx = \begin{cases} 0 & \text{for } i \neq k \\ >0 & \text{for } i = k \end{cases} \tag{11.70}$$

Let us consider some examples of orthogonal systems.

(1) The trigonometric system

$$\frac{1}{2},\quad \cos\frac{\pi x}{l},\quad \sin\frac{\pi x}{l},\ \ldots,\quad \cos\frac{k\pi x}{l},\quad \sin\frac{k\pi x}{l},\ \ldots \tag{11.71}$$

is orthogonal on $[-l, l]$.

(2) Each of the systems of functions

$$\text{(a)}\ \frac{1}{2},\quad \cos\frac{\pi x}{l},\quad \cos\frac{2\pi x}{l},\ \ldots,\quad \cos\frac{k\pi x}{l},\ \ldots$$

and (11.72)

$$\text{(b)}\ \sin\frac{\pi x}{l},\quad \sin\frac{2\pi x}{l},\ \ldots,\quad \sin\frac{k\pi x}{l},\ \ldots$$

is orthogonal on the interval $[0, l]$.

(3) The system of Legendre's* polynomials

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^n],\qquad n = 1, 2, \ldots,\quad P_0(x) = 1 \tag{11.73}$$

is orthogonal on the interval $[-1, 1]$ (see Appendix 1 to Chapter 11).

If a function $\varphi(x)$ is integrable on $[a, b]$ the nonnegative quantity

$$\|\varphi\| = \left(\int_a^b \varphi^2(x)\,dx\right)^{1/2} \tag{11.74}$$

is called the ***norm*** *of $\varphi(x)$ on $[a, b]$.* According to relations (11.70), the norms of all the functions of an orthogonal system are positive.

Using the notation (11.74) we can rewrite relations (11.70) as

$$\int_a^b \varphi_i(x)\,\varphi_k(x)\,dx = \begin{cases} 0 & \text{for } i \neq k \\ \|\varphi_k\|^2 & \text{for } i = k \end{cases} \tag{11.75}$$

---

* Legendre, Adrien Marie (1752-1833), a noted French mathematician.

We now consider the norms of the functions of some orthogonal systems.

(1) By definition (11.74) and relations (11.13$_2$), the norms of the functions forming trigonometric system (11.71) are

$$\left\|\frac{1}{2}\right\|=\sqrt{\frac{l}{2}}\qquad\left\|\cos\frac{k\pi x}{l}\right\|=\sqrt{l},\quad\left\|\sin\frac{k\pi x}{l}\right\|=\sqrt{l},\qquad k=1,\ 2,\ \ldots \tag{11.76}$$

(2) It can be easily verified that the norms of the functions of systems (11.72) considered on the interval $[0, l]$ are equal to

$$\text{(a)}\quad \left\|\frac{1}{2}\right\|=\frac{\sqrt{l}}{2},\quad\left\|\cos\frac{k\pi x}{l}\right\|=\sqrt{\frac{l}{2}},\qquad k=1,\ 2,\ \ldots$$

and (11.77)

$$\text{(b)}\quad \left\|\sin\frac{k\pi x}{l}\right\|=\sqrt{\frac{l}{2}},\qquad k=1,\ 2,\ \ldots$$

(3) The norms of Legendre's polynomials on the interval $[-1, 1]$ are

$$\|P_n(x)\|=\left(\int_{-1}^{+1}P_n^2(x)\,dx\right)^{1/2}=\sqrt{\frac{2}{2n+1}} \tag{11.78}$$

(the calculations connected with computing $\|P_n(x)\|$ are given in Appendix 1 to Chapter 11).

2. Fourier Coefficients and Fourier Series of a Function $f(x)$ with Respect to an Orthogonal System. Let a function $f(x)$ be integrable over $[a, b]$ and let the equality

$$f(x)=\sum_{k=1}^{+\infty}a_k\varphi_k(x) \tag{11.79}$$

hold where a_k are constant numbers and $\varphi_k(x)$ are the functions of an orthogonal system of type (11.69) on the interval $[a, b]$. The coefficients a_k can be easily expressed in terms of $f(x)$ if we multiply equality (11.79) by $\varphi_n(x)$, $n=1, 2, \ldots$, and integrate the relation thus obtained term-by-term over $[a, b]$.* Indeed, multiplying equality (11.79) by $\varphi_n(x)$ and integrating with respect to x from a to b we obtain the relation

$$\int_a^b f(x)\varphi_n(x)\,dx=\sum_{k=1}^{+\infty}a_k\int_a^b\varphi_k(x)\,\varphi_n(x)\,dx,\qquad n=1,\ 2,\ \ldots \tag{11.80}$$

* For termwise integration to be permissible it is sufficient that series (11.79) converge uniformly or in the mean to its sum on the interval $[a, b]$.

All the integrals on the right-hand side of equality (11.80) are equal to zero for $k \neq n$ (see orthogonality relations (11.70)). Consequently, we have

$$\int_a^b f(x)\varphi_n(x)\,dx = a_n \int_a^b \varphi_n^2(x)\,dx = a_n \|\varphi_n\|^2, \qquad n = 1, 2, \ldots$$

whence

$$a_n = \frac{1}{\|\varphi_n\|^2} \int_a^b f(x)\varphi_n(x)\,dx, \qquad n = 1, 2, \ldots \tag{11.81}$$

The numbers a_n determined by formulas (11.81) are called **Fourier's coefficients of the function $f(x)$ with respect to orthogonal system** (11.69) and series (11.79) whose coefficients are specified by formulas (11.81) is referred to as the **Fourier series of the function $f(x)$ with respect to orthogonal system** (11.69).

For the numbers a_k determined by formulas (11.81) to exist it is sufficient that the function $f(x)$ be integrable on $[a, b]$ (because the integrability of the functions $\varphi_k(x)$ is one of the conditions of the definition of an orthogonal system). Thus, with every function $f(x)$ integrable on $[a, b]$, we can associate its Fourier series with respect to system (11.69) orthogonal on $[a, b]$:

$$f(x) \sim \sum_{k=1}^{+\infty} a_k\varphi_k(x) \tag{11.82}$$

where the coefficients of series (11.82) are determined by formulas (11.81).

The conditions under which a given function $f(x)$ can be expanded into series (11.82) (i.e. the sign $\sim$ can be replaced by the sign of equality) depend on the properties of the orthogonal system $\{\varphi_k(x)\}$. In the case when we expand a given function with respect to the trigonometric system the conditions of the convergence theorem proved in § 2, Sec. 4 are sufficient for such an expansion to be valid. Similar convergence theorems for other orthogonal systems involve special investigation which we shall not present here.

3. Least Square Deviation. Bessel's Inequality. Consider an arbitrary fixed number of functions belonging to system (11.69) orthogonal on $[a, b]$:

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x) \tag{11.83}$$

A linear combination

$$\sum_{k=1}^{n} \alpha_k\varphi_k(x) \tag{11.84}$$

of functions (11.83) with arbitrary numerical coefficients α_k, $k = 1, 2, \ldots, n$, is called a polynomial of order n with respect orthogonal system (11.69).

Let $f(x)$ be an integrable function on $[a, b]$ and let it be necessar to solve the following problem.

It is required to determine the values of the coefficients α_k, $k = 1, 2, \ldots, n$, for which the corresponding polynomial of for: (11.84) has the *least mean square deviation* (see Chapter 8, § 6. Sec. *from the function $f(x)$ on the interval* $[a, b]$. Thus, we must fin the values of the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ for which the quantit

$$\rho^2\left(f, \sum_{k=1}^{n} \alpha_k\varphi_k\right) = \int_a^b \left[f(x) - \sum_{k=1}^{n} \alpha_k\varphi_k(x)\right]^2 dx = \left\| f - \sum_{k=1}^{n} \alpha_k\varphi_k \right\|^2 \tag{11.85}$$

achieves its absolute minimum.

Opening the square brackets we obtain

$$\rho^2\left(f, \sum_{k=1}^{n} \alpha_k\varphi_k\right) = \int_a^b f^2(x)\,dx - 2\sum_{k=1}^{n} \alpha_k \int_a^b f(x)\varphi_k(x)\,dx +$$

$$+ \int_a^b \left(\sum_{k=1}^{n} \alpha_k\varphi_k(x)\right)^2 dx = \int_a^b f^2(x)\,dx - 2\sum_{k=1}^{n} \alpha_k a_k \|\varphi_k\|^2 + \sum_{k=1}^{n} \alpha_k^2 \|\varphi_k\|^2 \tag{11.86}$$

because, by (11.81), we have $\int_a^b f(x)\varphi_k(x)\,dx = a_k\|\varphi_k\|^2$ and

$$\int_a^b \varphi_i(x)\varphi_k(x)\,dx = 0 \quad \text{for } i \neq k \quad \text{and} \quad \int_a^b \varphi_k^2(x)\,dx = \|\varphi_k\|^2$$

Completing the square we derive from (11.86) the relation

$$\rho^2\left(f, \sum_{k=1}^{n} \alpha_k\varphi_k\right) = \int_a^b f^2(x)\,dx - \sum_{k=1}^{n} a_k^2 \|\varphi_k\|^2 + \sum_{k=1}^{n} (\alpha_k - a_k)^2 \|\varphi_k\|^2 \tag{11.87}$$

Among the terms on the right-hand side of (11.87), only the last sum depends on the coefficients α_k. This sum being nonnegative, its greatest lower bound (equal to zero) is attained for $\alpha_k = a_k$. In this case the mean square deviation $\rho^2\left(f, \sum_{k=1}^{n} \alpha_k\varphi_k\right)$ achieves

its absolute minimum which is equal to

$$\min \rho^2\left(f, \sum_{k=1}^{n} \alpha_k\varphi_k\right) = \int_a^b \left[f(x) - \sum_{k=1}^{n} \alpha_k\varphi_k(x)\right]^2 dx =$$

$$= \int_a^b f^2(x)\,dx - \sum_{k=1}^{n} a_k^2 \|\varphi_k\|^2 \qquad (11.88)$$

The polynomial

$$\sum_{k=1}^{n} a_k\varphi_k(x) \qquad (11.89)$$

whose coefficients a_k are the Fourier coefficients of the function $f(x)$ with respect to the given orthogonal system $\{\varphi_k(x)\}$ is called **the Fourier polynomial of the function $f(x)$ with respect to the orthogonal system $\{\varphi_k(x)\}$.**

Thus, among all the polynomials of order n (where n is arbitrary but fixed) corresponding to the given system $\{\varphi_k(x)\}$ orthogonal on $[a, b]$, the Fourier polynomial of the function $f(x)$ with respect to the system $\{\varphi_k(x)\}$ has the least square deviation from the function $f(x)$ on the interval $[a, b]$.

Equality (11.88) expressing the least square deviation of the Fourier polynomial of a function $f(x)$ from this function is known as **Bessel's* identity.**

Noting that the left-hand side of equality (11.88) is nonnegative we obtain the inequality

$$\sum_{k=1}^{n} a_k^2 \|\varphi_k\|^2 \leqslant \int_a^b f^2(x)\,dx \qquad (11.90)$$

which holds for all $n \geqslant 1$ because its right-hand side is independent of n. The sum on the left-hand side of inequality (11.90) is non-decreasing as $n \to +\infty$ and therefore, since it is bounded from above by the integral $\int_a^b f^2(x)\,dx$, it tends to a finite limit for $n \to +\infty$. Passing to the limit in inequality (11.90) as $n \to +\infty$ we arrive at the so-called **Bessel's inequality:**

$$\sum_{k=1}^{+\infty} a_k^2 \|\varphi_k\|^2 \leqslant \int_a^b f^2(x)\,dx \qquad (11.91)$$

---

* Bessel, Friedrich Wilhelm (1784-1846), a German astronomer and mathematician.

By relations (11.76), in the case of the trigonometric system inequality (11.91) turns into

$$\frac{a_0^2}{2}+\sum_{k=1}^{+\infty}(a_k^2+b_k^2)\leqslant\frac{1}{l}\int_{-l}^{l}f^2(x)\,dx \tag{11.92}$$

It should be noted that we have established Bessel's inequality (11.91) for any function $f(x)$ which is integrable in the ordinary sense on the interval $[a, b]$ but it can be generalized to a wider class of functions.

We say that a function $f(x)$ is **square-integrable** or **square-summable (quadratically integrable or summable)** if the integrals

$$\int_a^b f(x)\,dx \quad\text{and}\quad \int_a^b f^2(x)\,dx \tag{11.93}$$

exist as proper or improper integrals.

It turns out that Bessel's inequality (11.91) and, consequently, inequality (11.92) remain valid for any function $f(x)$ square-integrable on $[a, b]$.

Moreover, Bessel's inequality (11.91) is also valid in the case when the functions $\varphi_k(x)$ of the orthogonal system are square-integrable on $[a, b]$. In fact, the convergence of the integrals $\int_a^b f^2(x)\,dx$ and $\int_a^b \varphi_k^2(x)\,dx$ and the evident inequalities

$$|f(x)\varphi_k(x)|\leqslant\frac{f^2(x)+\varphi_k^2(x)}{2} \quad\text{and}\quad |\varphi_i(x)\varphi_k(x)|\leqslant\frac{\varphi_i^2(x)+\varphi_k^2(x)}{2}\,^*$$

imply (by the comparison test) that the integrals

$$\int_a^b f(x)\varphi_k(x)\,dx \quad\text{and}\quad \int_a^b \varphi_i(x)\varphi_k(x)\,dx$$

are absolutely convergent. But when deducing Bessel's inequality we deal only with such integrals.**

Using the notion of a square-integrable function and introducing the concept of *orthogonality with weight function* (see Appendix 2 to Chapter 11) we can derive the *generalized Bessel's inequality* which holds for a still wider class of functions.

---

* In fact, we have $0\leqslant(|f(x)|-|\varphi_k(x)|)^2=f^2(x)-2|f(x)\varphi_k(x)|+\varphi_k^2(x)$ whence $|f(x)\varphi_k(x)|\leqslant\frac{1}{2}[f^2(x)+\varphi_k^2(x)]$.

** On some further generalizations, see Appendix 2 to Chapter 11.

Bessel's inequality for the trigonometric system (see inequality (11.92)) implies that the series $\frac{a_0^2}{2}+\sum_{k=1}^{+\infty}(a_k^2+b_k^2)$ is convergent and, consequently, we have

$$\lim_{k\to+\infty} a_k=\frac{1}{l}\lim_{k\to+\infty}\int_{-l}^{l} f(x)\cos\frac{k\pi x}{l}\,dx=0$$

and (11.94)

$$\lim_{k\to+\infty} b_k=\frac{1}{l}\lim_{k\to+\infty}\int_{-l}^{l} f(x)\sin\frac{k\pi x}{l}\,dx=0$$

Relations (11.94) are a special case of the more general relations

$$\lim_{\alpha\to+\infty}\int_{-l}^{l} f(x)\cos\alpha x\,dx=0 \quad\text{and}\quad \lim_{\alpha\to+\infty}\int_{-l}^{l} f(x)\sin\alpha x\,dx=0 \qquad (11.95)$$

which, according to Note 1 after the key lemma proved in § 2, are valid for any absolutely integrable function $f(x)$.

§ 4. SPEED OF CONVERGENCE OF FOURIER SERIES. ACCELERATION OF CONVERGENCE OF FOURIER SERIES

We shall first consider the conditions guaranteeing uniform convergence of a Fourier series and establish the relationship between the degree of smoothness of a function $f(x)$ and the rate at which the coefficients a_k and b_k of its Fourier series decrease as $k\to+\infty$. This will enable us to estimate the speed of convergence of the series.

1. Conditions for Uniform Convergence of Fourier Series. We shall prove that Fourier's trigonometric series of a continuous and piecewise smooth function $f(x)$ satisfying an additional necessary condition is uniformly convergent. We remind the reader that a function $f(x)$ is said to be continuous and piecewise smooth on an interval $[-l, l]$ if it is continuous and its derivative $f'(x)$ is piecewise continuous on this interval.

Theorem 11.2. If a continuous and piecewise smooth function $f(x)$ defined on an interval $[-l, l]$ assumes equal values at the end points of the interval, i.e. $f(-l)=f(l)$, its Fourier series

$$S(x)=\frac{a_0}{2}+\sum_{k=1}^{+\infty}\left(a_k\cos\frac{k\pi x}{l}+b_k\sin\frac{k\pi x}{l}\right) \qquad (11.96)$$

in which

$$a_0 = \frac{1}{l}\int_{-l}^{l} f(\xi)\,d\xi, \quad a_k = \frac{1}{l}\int_{-l}^{l} f(\xi)\cos\frac{n\pi\xi}{l}\,d\xi,$$

$$b_k = \frac{1}{l}\int_{-l}^{l} f(\xi)\sin\frac{n\pi\xi}{l}\,d\xi, \quad k = 1, 2, \ldots \tag{11.97}$$

is uniformly convergent on the interval $[-l, l]$ *and* $S(x) = f(x)$ *at each point of this interval.*

Proof. To establish the uniform convergence of series (11.96) it is sufficient to prove that the dominant number series

$$\frac{|a_0|}{2} + \sum_{k=1}^{+\infty} (|a_k| + |b_k|) \tag{11.98}$$

or, which is the same, the series

$$\sum_{k=1}^{+\infty} (|a_k| + |b_k|) \tag{11.99}$$

is convergent. Indeed, this will imply, by Weierstrass' test (see § 1, Sec. 2 in Chapter 8), that series (11.96) is uniformly convergent on the whole x-axis since we have

$$\left| a_k \cos\frac{k\pi x}{l} \right| \leqslant |a_k| \quad \text{and} \quad \left| b_k \sin\frac{k\pi x}{l} \right| \leqslant b_k, \quad k = 1, 2, \ldots$$

for all x, $-\infty < x < +\infty$. Let us denote the Fourier coefficients of the derivative $f'(x)$ by a_k' and b_k', i.e.

$$a_k' = \frac{1}{l}\int_{-l}^{l} f'(x)\cos\frac{k\pi x}{l}\,dx, \quad b_k' = \frac{1}{l}\int_{-l}^{l} f'(x)\sin\frac{k\pi x}{l}\,dx$$

Integrating by parts in the formulas expressing the Fourier coefficients of the function $f(x)$ we obtain

$$a_k = \frac{1}{l}\int_{-l}^{l} f(x)\cos\frac{k\pi x}{l}\,dx =$$

$$= \frac{1}{l}\,\frac{l}{k\pi} f(x)\sin\frac{k\pi x}{l}\Big|_{x=-l}^{x=l} - \frac{1}{l}\,\frac{l}{k\pi}\int_{-l}^{l} f'(x)\sin\frac{k\pi x}{l}\,dx = -\frac{l}{\pi}\,\frac{b_k'}{k}$$

$\left(\text{since } \sin\frac{k\pi x}{l} \text{ turns into zero for } x = \pm l\right)$ and

$$b_k = \frac{1}{l}\int_{-l}^{l} f(x)\sin\frac{k\pi x}{l}\,dx = -\frac{1}{l}\,\frac{l}{k\pi} f(x)\cos\frac{k\pi x}{l}\Big|_{x=-l}^{x=l} +$$

$$+ \frac{1}{l}\,\frac{l}{k\pi}\int_{-l}^{l} f'(x)\cos\frac{k\pi x}{l}\,dx = \frac{l}{\pi}\,\frac{a_k'}{k}$$

because, in view of the condition $f(l)=f(-l)$, we have

$$\cos\frac{k\pi x}{l} f(x)\Big|_{x=-l}^{x=l}=f(l)(-1)^k-f(-l)(-1)^k=$$
$$=(-1)^k[f(l)-f(-l)]=0$$

Therefore

$$|a_k|+|b_k|\leqslant\frac{l}{\pi}\left(\frac{|a'_k|}{k}+\frac{|b'_k|}{k}\right),\quad k=1, 2, \ldots \tag{11.100}$$

But, by the conditions of the theorem, the function $f'(x)$ is piecewise continuous on the interval $[-l, l]$ and hence it is integrable on this interval and Bessel's inequality

$$\frac{a_0'^2}{2}+\sum_{k=1}^{+\infty}(a_k'^2+b_k'^2)\leqslant\frac{1}{l}\int_{-l}^{l}f'^2(x)\,dx$$

is fulfilled for $f'(x)$. Consequently, the number series

$$\frac{a_0'^2}{2}+\sum_{k=1}^{+\infty}(a_k'^2+b_k'^2) \tag{11.101}$$

converges. Furthermore, the inequalities

$$0\leqslant\left(|a'_k|-\frac{1}{k}\right)^2=a'^2_k-2\frac{|a'_k|}{k}+\frac{1}{k^2}$$

and

$$0\leqslant\left(|b'_k|-\frac{1}{k}\right)^2=b_k'^2-2\frac{|b'_k|}{k}+\frac{1}{k^2}$$

imply the relations

$$\frac{|a'_k|}{k}\leqslant\frac{1}{2}\left(a_k'^2+\frac{1}{k^2}\right)\quad\text{and}\quad\frac{|b'_k|}{k}\leqslant\frac{1}{2}\left(b_k'^2+\frac{1}{k^2}\right)$$

Thus, we have

$$\frac{|a'_k|}{k}+\frac{|b'_k|}{k}\leqslant\frac{1}{2}(a_k'^2+b_k'^2)+\frac{1}{k^2} \tag{11.102}$$

The series $\sum\limits_{k=1}^{+\infty}\frac{1}{k^2}$ is convergent (this can easily be established on the basis of Cauchy's integral test), and therefore, since series (11.101) is also convergent, we conclude that the series $\sum\limits_{k=1}^{+\infty}\left\{\frac{1}{2}(a_k'^2+b_k'^2)+\frac{1}{k^2}\right\}$ converges. Consequently, by the compari-

son test and inequalities (11.100) and (11.102), the series

$$\frac{|a_0|}{2}+\sum_{k=1}^{+\infty}(|a_k|+|b_k|)$$

(which is dominant for the Fourier series of the function $f(x)$) also converges. It follows that Fourier's trigonometric series of the function $f(x)$ is uniformly convergent to its sum $S(x)$ on the entire x-axis. The validity of the equality $S(x) = f(x)$ on the interval $[-l, l]$ is implied by the convergence theorem proved in Sec. 4 of § 2 because the conditions of this theorem are fulfilled here. Thus, the theorem has been proved.

Note. The condition that the values of the function $f(x)$ at the end points of the interval $[-l, l]$ are equal is necessary for Fourier's series (11.96) of the function $f(x)$ to be convergent to $f(x)$ at the end points. In fact, if the sum of the series satisfies the equalities

$$S(-l) = f(-l), \quad S(l) = f(l) \tag{11.103}$$

the relation

$$S(-l) = S(l) \tag{11.104}$$

(which is a consequence of the periodicity of the sum of series (11.96) whose all terms are periodic with period $2l$) implies the equality

$$f(-l) = f(l) \tag{11.105}$$

Therefore, for Fourier series (11.96) of the function $f(x)$ to be uniformly convergent to $f(x)$ on the closed interval $[-l, l]$ it is necessary that equality (11.105) hold.

If the values of a continuous and piecewise smooth function $f(x)$, which is defined on an interval $[-l, l]$ and assumes equal values at its end points, are changed at a finite number of points this results in a discontinuous function whose values at the end points of the interval may no longer coincide. But the Fourier coefficients of the modified function remain the same since the integrals expressing these coefficients do not change. Consequently, by the estimates obtained in the proof of Theorem 11.2, the series $\sum\limits_{k=1}^{+\infty}(|a_k|+|b_k|)$ is convergent and thus the Fourier series of the modified piecewise continuous and piecewise smooth function is also uniformly convergent on the interval $[-l, l]$ and its sum is equal to the original function but not to the modified function. But it can be proved that for the Fourier series of a piecewise continuous and piecewise smooth function $f(x)$ defined on an interval $[-l, l]$ to be uniformly convergent on that interval, it is necessary that all the discontinuities of the function be removable and that the equality of the limiting values at the end points of the interval hold, i.e. $f(-l+0)$

$= f(l-0)$. This can easily be proved on the basis of the theorems on continuity of the sum of a uniformly convergent series and the possibility of termwise passage to the limit in such a series (see § 2 of Chapter 8).

Theorem 11.2 can be stated in a different manner. First of all note that if a function $f(x)$ is continuous on an interval $[-l, l]$ and takes on equal values at its end points, the function obtained when $f(x)$ is periodically extended with period $2l$ is continuous throughout the x-axis.

Next, let us agree that a function $f(x)$ will be called *piecewise smooth on the whole x-axis* if it is piecewise smooth on every finite interval of the x-axis. Then Theorem 11.2 can be restated as follows:

If a periodic function $f(x)$ with period $2l$ is continuous and piecewise smooth on the entire x-axis its Fourier series (11.96) *converges uniformly to this function on the x-axis.*

2. Connection Between the Degree of Smoothness of a Function and the Speed of Convergence of Its Fourier Series. The investigation of the connection between the properties of a function $f(x)$ and the speed of convergence of its Fourier series is important for applications of Fourier's series to some problems of mathematical physics in which the sum of such a series is replaced by its nth partial sum. It is also important in view of the possibility of term-by-term differentiation of such series etc.

Theorem 11.3. If a function $f(x)$ and its derivatives $f'(x)$, ..., $f^{(m)}(x)$ $(m \geqslant 0)$ are continuous on an interval $[-l, l]$ and assume equal values at its end points, that is

$$f(-l) = f(l),\ f'(-l) = f'(l),\ \ldots,\ f^{(m)}(-l) = f^{(m)}(l) \tag{11.106}$$

and the derivative $f^{(m+1)}(x)$ is piecewise continuous on the interval $[-l, l]$, the Fourier coefficients

$$a_k = \frac{1}{l}\int_{-l}^{l} f(\xi)\cos\frac{k\pi\xi}{l}\,d\xi \quad \textit{and} \quad b_k = \frac{1}{l}\int_{-l}^{l} f(\xi)\sin\frac{k\pi\xi}{l}\,d\xi$$

of the function $f(x)$ satisfy the relations

$$a_k = o\left(\frac{1}{k^{m+1}}\right) \textit{ and } b_k = o\left(\frac{1}{k^{m+1}}\right) \quad \textit{for } k\to\infty, \tag{11.107}$$

and the series

$$\sum_{k=1}^{+\infty} k^{\nu}(|a_k| + |b_k|), \quad \nu = 0, 1, 2, \ldots, m \tag{11.108}$$

are convergent. *

Proof. Integrating by parts, as was done in the proof of Theorem 11.1, we obtain

$$a_k=\frac{1}{l}\int_{-l}^{l} f(\xi)\cos\frac{k\pi\xi}{l}\,d\xi=$$

$$=\frac{1}{l}\,\frac{l}{k\pi}\,f(\xi)\sin\frac{k\pi\xi}{l}\Big|_{\xi=-l}^{\xi=l}-\frac{1}{k\pi}\int_{-l}^{l} f'(\xi)\sin\frac{k\pi\xi}{l}\,d\xi=$$

$$=-\frac{1}{k\pi}\int_{-l}^{l} f'(\xi)\sin\frac{k\pi\xi}{l}\,d\xi=$$

$$=\frac{l}{k^2\pi^2}\,f'(\xi)\cos\frac{k\pi\xi}{l}\Big|_{\xi=-l}^{\xi=l}-\frac{l}{k^2\pi^2}\int_{-l}^{l} f''(\xi)\cos\frac{k\pi\xi}{l}\,d\xi=$$

$$=-\frac{l}{k^2\pi^2}\int_{-l}^{l} f''(\xi)\cos\frac{k\pi\xi}{l}\,d\xi=\ldots$$

$$\ldots=\pm\frac{l^{m+1}}{k^{m+1}\pi^{m+1}}\,\frac{1}{l}\int_{-l}^{l} f^{(m+1)}(\xi)\begin{Bmatrix}\cos\dfrac{k\pi\xi}{l}\\ \sin\dfrac{k\pi\xi}{l}\end{Bmatrix}d\xi \qquad (11.109)$$

In these calculations we take into account that (1) condition (11.10) of the theorem is fulfilled, (2) the cosine is an even function and (3) the sine of the argument $\frac{k\pi\xi}{l}$, $k=1, 2, \ldots$, turns into zero at the end points of the interval $[-l, l]$ and therefore we have the relations

$$f^{(s)}(\xi)\cos\frac{k\pi\xi}{l}\Big|_{\xi=-l}^{\xi=l}=0 \quad\text{and}\quad f^{(s)}(\xi)\sin\frac{k\pi\xi}{l}\Big|_{\xi=-l}^{\xi=l}=0$$
$$\text{for}\quad 0\leqslant s\leqslant m$$

The curly brackets under the sign of integration in (11.109) mean that, depending on the number of times the integration by parts has been performed, the factor by which the function $f^{(m+1)}(\xi)$ is multiplied is either $\cos\frac{k\pi\xi}{l}$ or $\sin\frac{k\pi\xi}{l}$.

---

* Relations (11.107) mean that

$$\lim_{k\to+\infty}\frac{a_k}{\dfrac{1}{k^{m+1}}}=0 \quad\text{and}\quad \lim_{k\to+\infty}\frac{b_k}{\dfrac{1}{k^{m+1}}}=0$$

In just the same way we obtain

$$b_k = \pm \frac{l^{m+1}}{k^{m+1}\pi^{m+1}} \frac{1}{l} \int_{-l}^{l} f^{(m+1)}(\xi) \begin{Bmatrix} \sin \frac{k\pi\xi}{l} \\ \cos \frac{k\pi\xi}{l} \end{Bmatrix} d\xi \qquad (11.110)$$

From (11.109) and (11.110) we derive

$$|a_k| + |b_k| = \frac{l^{m+1}}{\pi^{m+1}} \left(\frac{|a_k^{(m+1)}|}{k^{m+1}} + \frac{|b_k^{(m+1)}|}{k^{m+1}} \right) \qquad (11.111)$$

where $a_k^{(m+1)}$ and $b_k^{(m+1)}$ are the Fourier coefficients of $f^{(m+1)}(x)$.

Since $a_k^{(m+1)}$ and $b_k^{(m+1)}$ tend to zero as $k \to +\infty$, it follows from (11.111) that

$$a_k = o\left(\frac{1}{k^{m+1}}\right), \quad b_k = o\left(\frac{1}{k^{m+1}}\right) \quad \text{for } y \to +\infty \qquad (11.112)$$

Relation (11.111) implies that

$$k^m (|a_k| + |b_k|) \leqslant \frac{l^{m+1}}{\pi^{m+1}} \left(\frac{|a_k^{(m+1)}|}{k} + \frac{|b_k^{(m+1)}|}{k} \right) \leqslant$$

$$\leqslant \frac{l^{m+1}}{2\pi^{m+1}} \left\{ |a_k^{(m+1)}|^2 + |b_k^{(m+1)}|^2 + \frac{2}{k^2} \right\} \qquad (11.113)$$

because

$$\frac{|a_k^{(m+1)}|}{k} \leqslant \frac{1}{2} \left(|a_k^{(m+1)}|^2 + \frac{1}{k^2} \right), \quad \frac{|b_k^{(m+1)}|}{k} \leqslant \frac{1}{2} \left(|b_k^{(m+1)}|^2 + \frac{1}{k^2} \right)$$

Therefore, by Bessel's inequality

$$\frac{|a_0^{(m+1)}|^2}{2} + \sum_{k=1}^{+\infty} (|a_k^{(m+1)}|^2 + |b_k^{(m+1)}|^2) \leqslant \frac{1}{l} \int_{-l}^{l} |f^{(m+1)}(x)|^2 \, dx$$

and by the convergence of the series $\sum\limits_{k=1}^{+\infty} \frac{1}{k^2}$ we conclude that the series $\sum\limits_{k=1}^{+\infty} k^m (|a_k| + |b_k|)$ is convergent. It follows that all the series (11.108) are convergent and thus the theorem has been proved.

Note 1. To apply the above theorem to the expansion of a function $f(x)$ which is defined only on an interval $[0, l]$ into a series with respect to the system of functions $\sin \frac{k\pi x}{l}$, $k = 1, 2, \ldots$, we must verify the validity of the conditions of Theorem 11.3 for the function $F(x)$ defined on the entire interval $[-l, l]$ which is obtained when $f(x)$ is extended as an odd function. In particular, for the function $F(x)$ to be continuous at the point $x = 0$ it is neces-

sary that the equality $f(0) = 0$ hold because, if otherwise, the extended function would have a discontinuity at $x = 0$. Analogously, at the point $x = l$ we must have $f(l) = 0$ because the extended function should satisfy the equality $F(-l) = F(l)$. The derivative of an odd function being even, the condition $F'(-l) = F'(l)$ automatically holds for the derivative of the function $F(x)$.

Generally, for the derivatives of the function $F(x)$ to be continuous on the interval $[-l, l]$ and to have equal values at its end points we should impose the conditions

$$f^{(\nu)}(0) = f^{(\nu)}(l) = 0 \quad \text{for} \quad \nu = 0, 2, 4, \ldots$$

Then the corresponding conditions of Theorem 11.3 automatically hold for the derivatives of $F(x)$ of odd orders.

In particular, for the series

$$\sum_{k=1}^{+\infty} k^{\nu} |b_k|, \quad \nu = 0, 1, 2$$

(where $b_k = \frac{2}{l} \int\limits_0^l f(\xi) \sin \frac{k\pi\xi}{l} \, d\xi$ are the Fourier coefficients of a function $f(x)$ defined on the interval $[0, l]$) to be convergent it is sufficient that the function $f(x)$ satisfy the following conditions:

(1) the function $f(x)$ and its derivatives $f'(x)$ and $f''(x)$ are continuous and the derivative $f'''(x)$ is piecewise continuous on the interval $[0, l]$;

(2) $f(0) = f(l)$ and $f''(0) = f''(l) = 0$.

Note 2. If a function $f(x)$ satisfying the conditions of Theorem 11.3 on an interval $[-l, l]$ is periodically extended with period $2l$ from the interval $[-l, l]$ to the whole x-axis, the extended function and its derivatives up to the order m inclusive are continuous and periodic with period $2l$ for all x, $-\infty < x < +\infty$. Therefore Theorem 11.3 can be rephrased as follows:

If a periodic function $f(x)$ with period $2l$ and its derivatives up to the mth order inclusive ($m \geqslant 0$) are continuous for $-\infty < x < +\infty$ and the $(m+1)$th derivative $f^{(m+1)}(x)$ is piecewise continuous, the Fourier coefficients a_k and b_k of this function with respect to trigonometric system (11.12) satisfy the following conditions:

(1) $\quad a_k = o\left(\frac{1}{k^{m+1}}\right)$ *and* $b_k = o\left(\frac{1}{k^{m+1}}\right)$ *for* $k \to +\infty$

(2) *the series* $\sum\limits_{k=1}^{+\infty} k^{\nu} (|a_k| + |b_k|), \quad \nu = 0, 1, \ldots, m,$

are convergent.

Hence, this theorem establishes a relationship between the degree of smoothness of a periodic function and the speed of convergence of its Fourier series.

Note 3. If the conditions of Theorem 11.3 are fulfilled for $m > 0$ the Fourier series of the function $f(x)$ can be differentiated term-by-term at least m times, that is we have the relations

$$f^{(s)}(x) = \sum_{k=1}^{+\infty} \left(a_k \cos\frac{k\pi x}{l} + b_k \sin\frac{k\pi x}{l}\right)^{(s)} \tag{11.114}$$

$$\text{for} \quad 1 \leqslant s \leqslant m, \quad -l \leqslant x \leqslant l$$

since the dominant series

$$\frac{\pi^s}{l^s} \sum_{k=1}^{+\infty} k^s (|a_k| + |b_k|), \quad 1 \leqslant s \leqslant m$$

are convergent.

Note 4. The proof of Theorem 11.3 enables us to estimate the speed of convergence of a Fourier series, i.e. to obtain an estimation for the error arising when the sum of the Fourier series is replaced by its partial sum. In fact, under the conditions of Theorem 11.3, we can apply relation (11.111), the Cauchy-Bunyakovsky inequality for sums, Bessel's inequality for Fourier's coefficients of the function $f^{(m+1)}(x)$ and the obvious inequality

$$\sum_{k=k_0+1}^{+\infty} \frac{1}{k^{2m+2}} \leqslant \int_{k_0}^{+\infty} \frac{dx}{x^{2m+2}}$$

which leads to the relation

$$\left|\sum_{k=k_0+1}^{+\infty} \left(a_k \cos\frac{k\pi x}{l} + b_k \sin\frac{k\pi x}{l}\right)\right| \leqslant \sum_{k=k_0+1}^{+\infty} (|a_k| + |b_k|) \leqslant$$

$$\leqslant \frac{l^{m+1}}{\pi^{m+1}} \sqrt{\sum_{k=k_0+1}^{+\infty} \frac{1}{k^{2m+2}}} \sqrt{2 \sum_{k=k_0+1}^{+\infty} \left(|a_k^{(m+1)}|^2 + |b_k^{(m+1)}|^2\right)} \leqslant$$

$$\leqslant \frac{l^{m+1}}{\pi^{m+1}} \left(\int_{k_0}^{+\infty} \frac{dx}{x^{2m+2}}\right)^{1/2} \left(\frac{2}{l} \int_{-l}^{l} |f^{(m+1)}(x)|^2 \, dx\right)^{1/2} =$$

$$= \frac{2l^{m+\frac{1}{2}}}{\pi^{m+1}(2m+1)^{1/2}} \left(\int_{-l}^{l} |f^{(m+1)}(x)|^2 \, dx\right)^{1/2} \frac{1}{k_0^{m+\frac{1}{2}}} = O\left(\frac{1}{k_0^{m+\frac{1}{2}}}\right)$$

3. Acceleration of Convergence of Fourier Series. A Fourier series with respect to the trigonometric system appearing in a concrete problem may turn out to converge so slowly that it is impossible to use it for practical purposes when its sum is unknown.

In this connection it is natural to pose the question whether it is possible to choose for a given slowly converging Fourier series

$$f(x)=\frac{a_0}{2}+\sum_{k=1}^{+\infty}\left(a_k\cos\frac{k\pi x}{l}+b_k\sin\frac{k\pi x}{l}\right),\quad -l\leqslant x\leqslant l \qquad (A)$$

a new slowly converging trigonometric series whose sum $\varphi(x)$ is known such that the series entering into the relation

$$f(x)=\varphi(x)+\sum_{k=1}^{+\infty}\left(\alpha_k\cos\frac{k\pi x}{l}+\beta_k\sin\frac{k\pi x}{l}\right) \qquad (B)$$

converges fast enough, that is such that its coefficients α_k and β_k tend to zero sufficiently fast as $k\to+\infty$.

When passing from representation (A) of the function $f(x)$ to representation (B), we say that the convergence of series (A) is *accelerated* (*improved*).

If some singularities of the function $f(x)$ (such as the limiting values of $f(x)$ and of the derivatives $f^{(s)}(x)$, $s=1, 2, \ldots, m$, at their points of discontinuity and for $x\to\pm l$ etc.) are known the convergence can be accelerated by means of a simple function $\varphi(x)$ (which is known) having the same singularities as $f(x)$.

For instance, suppose we know that $f(x)$ is continuously differentiable on the interval $[-l, l]$ and $\lim\limits_{x\to\pm l} f(x)=\pm l$. The values of $f(x)$ assumed at the end points of the interval $[-l, l]$ being unequal, series (A) converges nonuniformly on this interval. Let us put $\varphi(x)=x$. The latter function is also continuously differentiable on the interval $[-l, l]$ and possesses the same limiting values at its end points as $f(x)$. Therefore the function $f(x)-x$ is continuously differentiable on the interval $[-l, l]$ and its end point limiting values are equal. Consequently, the series entering into representation (B) of the function $f(x)$ for $\varphi(x)=x$ is uniformly convergent on the interval $[-l, l]$.

Now let us discuss another method of improving convergence of series (A) which is applicable when we have no additional information concerning its sum. This approach was suggested by A. N. Krylov*. The basic idea of the method is that we select those coefficients a_n and b_n of series (A) which contain the lowest powers of the quantity $\frac{1}{n}$ and try to compute the sum of the auxiliary series with these coefficients by using the tables of Fourier series expansions of various functions whose Fourier series are slowly convergent.

---

* Krylov, Aleksei Nikolayevich (1863-1945), a prominent Russian mathematician and engineer.

For example, let it be necessary to improve the convergence of the series

$$f(x) = \sum_{n=2}^{+\infty} (-1)^n \frac{n^3}{n^4-1} \sin nx, \quad -\pi < x < \pi \qquad (11.115)$$

We have

$$\frac{n^3}{n^4-1} = \frac{1}{n} + \frac{1}{n^5-n}$$

and therefore

$$f(x) = \sum_{n=2}^{+\infty} (-1)^n \frac{\sin nx}{n} + \sum_{n=2}^{+\infty} (-1)^n \frac{\sin nx}{n^5-n}, \quad -\pi < x < \pi$$

But we have already found (see Example 1 in § 2, Sec. 5) that

$$x = 2\sum_{n=1}^{+\infty} (-1)^{n+1} \frac{\sin nx}{n}, \quad -\pi < x < \pi$$

(to obtain the last expression we must put $l = \pi$ in the above-mentioned example). Hence, we can write

$$\sum_{n=2}^{+\infty} (-1)^n \frac{\sin nx}{n} = -\frac{x}{2} + \sin x, \quad -\pi < x < \pi$$

Consequently,

$$f(x) = -\frac{x}{2} + \sin x + \sum_{n=2}^{+\infty} (-1)^n \frac{\sin nx}{n^5-n}, \quad -\pi < x < \pi \qquad (11.116)$$

Series (11.116) converges much faster than the original series (11.115).

§ 5. UNIFORM APPROXIMATION OF CONTINUOUS FUNCTION BY TRIGONOMETRIC AND ALGEBRAIC POLYNOMIALS. WEIERSTRASS' APPROXIMATION THEOREMS

Let $\varepsilon > 0$ be an arbitrary fixed number. If the inequality

$$|\varphi(x) - \psi(x)| < \varepsilon$$

is fulfilled for all $x \in [a, b]$ simultaneously we say that *the function $\varphi(x)$ is uniformly approximated on the interval $[a, b]$ by the function $\psi(x)$ to an accuracy of ε.*

Theorem 11.4 (Weierstrass' Trigonometric Approximation Theorem). *If a function $f(x)$ is continuous on an interval $-l \leqslant x \leqslant l$ and assumes equal values at its end points, then*

for every $\varepsilon > 0$ *there is a* ***trigonometric polynomial***

$$T_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos \frac{k\pi x}{l} + b_k \sin \frac{k\pi x}{l}\right) \qquad (11.117)$$

which uniformly approximates the function $f(x)$ *on the interval* $-l \leqslant x \leqslant l$ *to an accuracy of* ε.

The proof of this theorem is based on the following

Lemma. *For every continuous function* $f(x)$ *defined on an interval* $a \leqslant x \leqslant b$ *and for every* $\varepsilon > 0$ *there exists a continuous piecewise smooth function* $g_\varepsilon(x)$ *defined on this interval such that*

$$|f(x) - g_\varepsilon(x)| < \frac{\varepsilon}{2} \quad \textit{for all} \quad x \in [a, b] \qquad (11.118)$$

and

$$g_\varepsilon(a) = f(a), \quad g_\varepsilon(b) = f(b) \qquad (11.119)$$

Proof of the lemma. Since the function $f(x)$ is continuous on the closed interval $[a, b]$ it is uniformly continuous on it, that is for every $\varepsilon > 0$ there is $\delta = \delta(\varepsilon) > 0$ such that the inequality

$$|f(x') - f(x'')| < \frac{\varepsilon}{2} \qquad (11.120)$$

is fulfilled for any x' and x'' belonging to the interval $[a, b]$ and satisfying the condition $|x' - x''| < \delta(\varepsilon)$. Therefore, if we break up the interval $[a, b]$ into subintervals $[x_i, x_{i+1}]$, $i = 0, 1, \ldots, n$, of lengths less than δ by means of points of division $x_0 = a < x_1 < \ldots < x_i < x_{i+1} < \ldots < x_{n+1} = b$, inequality (11.120) holds for any two points x' and x'' belonging to one and the same subinterval $[x_i, x_{i+1}]$.

Let us construct a continuous piecewise smooth function $y = g_\varepsilon(x)$ defined on the interval $[a, b]$ by putting $g_\varepsilon(x_i) = f(x_i)$ for $i = 0, 1, \ldots, n+1$ and $g_\varepsilon(x) = g_\varepsilon(x_i) + \frac{g_\varepsilon(x_{i+1}) - g_\varepsilon(x_i)}{x_{i+1} - x_i}(x - x_i)$ for $x_i \leqslant x \leqslant x_{i+1}$. This means that the function $y = g_\varepsilon(x)$ is linear on each subinterval $[x_i, x_{i+1}]$, $i = 0, 1, \ldots, n$, and its graph is a polygonal line inscribed in the graph of the function $y = f(x)$. According to the construction of $g_\varepsilon(x)$ we have

$$g_\varepsilon(a) = f(a), \quad g_\varepsilon(b) = f(b)$$

Now we can show that

$$|f(x') - g_\varepsilon(x')| < \frac{\varepsilon}{2}$$

for any $x' \in [a, b]$. Indeed, let, for instance, $x' \in [x_i, x_{i+1}]$. The function $g_\varepsilon(x)$ being linear on the subinterval $[x_i, x_{i+1}]$, the value $g_\varepsilon(x')$ lies between the values $g_\varepsilon(x_i) = f(x_i)$ and $g_\varepsilon(x_{i+1}) =$

$= f(x_{i+1})$. As is known, a continuous function assumes all intermediate values lying between its end-point values and therefore the function $f(x)$ takes on all intermediate values between $f(x_i)$ and $f(x_{i+1})$ on the interval $[x_i, x_{i+1}]$. Hence, there is $x'' \in [x_i, x_{i+1}]$ such that $f(x'') = g_\varepsilon(x')$. Consequently,

$$|f(x') - g_\varepsilon(x')| = |f(x') - f(x'')| < \frac{\varepsilon}{2}$$

since $x', x'' \in [x_i, x_{i+1}]$, which is what we set out to prove.

Proof of Theorem 11.4. By the hypothesis, the function $f(x)$ is continuous on the interval $[-l, l]$ and has equal values at its end points: $f(-l) = f(l)$. Let us take an arbitrary $\varepsilon > 0$. According to the above lemma, there is a continuous piecewise smooth function $g_\varepsilon(x)$ defined on $[-l, l]$ such that

$$|f(x) - g_\varepsilon(x)| < \frac{\varepsilon}{2} \text{ for all } x \in [-l, l] \tag{11.121}$$

and

$$g_\varepsilon(-l) = f(-l), \quad g_\varepsilon(l) = f(l)$$

which implies

$$g_\varepsilon(-l) = g_\varepsilon(l) \tag{11.122}$$

since

$$f(-l) = f(l)$$

By Theorem 11.2, Fourier's series of the function $g_\varepsilon(x)$ is uniformly convergent to $g_\varepsilon(x)$ on the interval $[-l, l]$. Hence, for a sufficiently large n, the nth partial sum $T_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n}\left(a_k \cos\frac{k\pi x}{l} + b_k \sin\frac{k\pi x}{l}\right)$ of this series satisfies the relation

$$|g_\varepsilon(x) - T_n(x)| < \frac{\varepsilon}{2} \text{ for all } x \in [-l, l] \tag{11.123}$$

From (11.121) and (11.123) we conclude that

$$|f(x) - T_n(x)| \leqslant |f(x) - g_\varepsilon(x)| + |g_\varepsilon(x) - T_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \tag{11.124}$$

for all $x \in [-l, l]$. The theorem has been proved.

Note. Taking a number sequence $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k, \ldots$ convergent to zero we can construct the corresponding sequence of trigonometric polynomials $T_{n_1}(x), T_{n_2}(x), \ldots$ uniformly convergent to the function $f(x)$ on the interval $[-l, l]$. But in the general case these trigonometric polynomials are not partial sums of one and the same trigonometric series. In fact, the polynomial $T_n(x)$ (corres-

ponding to a given $\varepsilon > 0$) which enters into inequality (11.124) is Fourier's polynomial of the auxiliary continuous piecewise smooth function $g_\varepsilon(x)$ which changes when ε varies and hence the coefficients of the polynomial $T_n(x)$ also vary. But it should be noted that the above fact is not a consequence of the specific method used for constructing the polynomials $T_n(x)$ in the proof of Theorem 11.4 because it can be shown that in the general case a continuous function $f(x)$ may not be the limit of a uniformly convergent sequence of partial sums of one and the same trigonometric series. Indeed, if $f(x)$ were the limit of a uniformly convergent sequence of partial sums of a trigonometric series

$$\frac{\alpha_0}{2}+\sum_{k=1}^{+\infty}\left(\alpha_k\cos\frac{k\pi x}{l}+\beta_k\sin\frac{k\pi x}{l}\right)$$

on the interval $[-l, l]$ this series would necessarily be the Fourier series of $f(x)$. But there are examples of continuous functions $f(x)$ defined on an interval $[-l, l]$ whose Fourier series diverge at a finite (and even at an infinite) number of points belonging to the interval $[-l, l]$. Such examples are rather difficult to construct and we shall not consider them here.*

Theorem 11.5. (Weierstrass' Polynomial Approximation Theorem). *If $f(x)$ is continuous on an interval $a \leqslant x \leqslant b$, then for every $\varepsilon > 0$ there exists an algebraic polynomial*

$$P_m(x) = A_0 + A_1x + A_2x^2 + \ldots + A_mx^m$$

such that

$$|f(x) - P_m(x)| < \varepsilon \tag{11.125}$$

on the interval $[a, b]$.

Proof. Take a sufficiently large $l > 0$ such that the interval $[a, b]$ is strictly contained within the interval $[-l, l]$. Let us construct a continuous function $F(x)$ defined on $[-l, l]$ by putting $F(x) = f(x)$ for $a \leqslant x \leqslant b$, $F(-l) = F(l) = 0$,

$$F(x)=\frac{f(a)}{a+l}(x+l) \text{ for } -l\leqslant x\leqslant a \text{ and } F(x)=\frac{f(b)}{l-b}(l-x)$$

for $b \leqslant x \leqslant l$ (i.e. $F(x)$ is a linear function on the intervals $[-l, a]$ and $[b, l]$). By the Weierstrass' trigonometric approximation theorem, for every $\varepsilon > 0$ there is a trigonometric polynomial

$$T_n(x)=\frac{a_0}{2}+\sum_{k=1}^{n}\left(a_k\cos\frac{k\pi x}{l}+b_k\sin\frac{k\pi x}{l}\right) \tag{11.126}$$

* E.g. see [1].

such that for all $-l \leqslant x \leqslant l$ we have

$$|F(x)-T_n(x)|<\frac{\varepsilon}{2} \tag{11.127}$$

Let us expand into Taylor's series the trigonometric functions entering into (11.127):

$$\cos\frac{k\pi x}{l}=1-\frac{k^2\pi^2}{2!l^2}x^2+\frac{k^4\pi^4}{4!l^4}x^4-\ldots \tag{11.128}$$

and

$$\sin\frac{k\pi x}{l}=\frac{k\pi}{l}x-\frac{k^3\pi^3}{3!l^3}x^3+\frac{k^5\pi^5}{5!l^5}x^5-\ldots \tag{11.129}$$

Power series (11.128) and (11.129) are absolutely convergent for all x, $-\infty<x<+\infty$, which can be, for instance, proved by applying D'Alembert's test. Substituting (11.128) and (11.129) into (11.126) we arrive at a power series

$$T_n(x)=A_0+A_1x+A_2x^2+\ldots+A_mx^m+\ldots \tag{11.130}$$

which is convergent for all x, $-\infty<x<+\infty$. Consequently, series (11.130) converges uniformly on every finite interval of the x-axis and, in particular, on the interval $[-l, l]$. Therefore, taking a sufficiently large m we obtain the mth partial sum

$$P_m(x)=A_0+A_1x+\ldots+A_mx^m$$

of series (11.130) such that

$$|T_n(x)-P_m(x)|<\frac{\varepsilon}{2} \text{ for all } x\in[-l, l] \tag{11.131}$$

From inequalities (11.127) and (11.131) it follows that

$$|F(x)-P_m(x)|\leqslant|F(x)-T_n(x)|+|T_n(x)-P_m(x)|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon \tag{11.132}$$

for all $x\in[-l, l]$ and, in particular, for all $x\in[a, b]$. But we have $F(x)=f(x)$ for all $x\in[a, b]$ and thus inequality (11.132) turns into the relation

$$|f(x)-P_m(x)|<\varepsilon \tag{11.133}$$

which is what we set out to prove.

Note. If we take a number sequence $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k, \ldots$ convergent to zero we can construct the corresponding sequence of algebraic polynomials $P_{m_1}(x)$, $P_{m_2}(x)$, ... uniformly convergent to $f(x)$ on $[a, b]$. But in the general case these polynomials may not be partial sums of one and the same power series which can be confirmed by an argument similar to the presented above in connection with trigonometric polynomials (see the note after Theorem 11.4).

Weierstrass' theorems do not provide effective means for practical construction of polynomials uniformly approximating a given continuous function to any preassigned degree of accuracy $\varepsilon > 0$.

The problem of constructing polynomials yielding the best approximation was posed and solved by Chebyshev*. Chebyshev's polynomials make it possible to construct effectively polynomials approximating a given continuous function.

We now denote by H_n the totality of all algebraic polynomials P_m of degree $m \leqslant n$. Let $P_m(x) \in H_n$ and let $f(x)$ be a continuous function on $[a, b]$. The number

$$E(f, P_m(x)) = \max_{a \leqslant x \leqslant b} |f(x) - P_m(x)|$$

is termed **the deviation of** $P_m(x)$ **from** $f(x)$ **on** $[a, b]$. The greatest lower bound $E_n(f)$ of the values of the quantity $E(f, P_m)$ when $P_m(x)$ runs through the whole set H_n is called the **least deviation**. Chebyshev proved the existence and uniqueness theorem for a polynomial of the best uniform approximation, that is a polynomial $P_m(x) \in H_n$ such that

$$E(f, P_m) = E_n(f)$$

and also developed some methods of constructing such polynomials. He found the so-called **polynomials of least deviation from zero**, also known as **Chebyshev's polynomials** (see [2], vol. 2, Chapter 4). In practical problems of constructing a polynomial uniformly approximating a continuous function $f(x)$ on an interval $[a, b]$ to an accuracy of ε it is important to obtain a polynomial whose degree is as low as possible. It is obvious that such a polynomial is the polynomial of the best approximation (for the function $f(x)$ on the interval $[a, b]$) belonging to the totality H_n with n such that

$$E_n(f) \leqslant \varepsilon < E_{n-1}(f)$$

§ 6. COMPLETE AND CLOSED ORTHOGONAL SYSTEMS

Let us denote by $Q[a, b]$ the class of functions containing all the piecewise continuous functions defined on $[a, b]$. We shall introduce the notions of complete and closed orthogonal systems for the functions belonging to the class $Q[a, b]$. Basic theorems concerning these notions will also be proved for this class of functions (Theorems 11.6-11.10). But it should be noted that this theory can be generalized to a considerably wider class of functions, namely to the square-integrable functions on $[a, b]$ and even to the functions

* Chebyshev, Pafnuty Lvovich (1821-1894), a famous Russian mathematician.

which are square-integrable with a weight function $p(x)$ on $[a, b]$ (see Appendix 2 to Chapter 11).

1. Complete Orthogonal System. An orthogonal system of functions

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots \tag{11.134}$$

on $[a, b]$ is said to be complete if for every function $f(x)$ belonging to $Q[a, b]$ its Fourier series with respect to orthogonal system (11.134)

$$\sum_{k=1}^{+\infty} c_k \varphi_k(x) \tag{11.135}$$

where

$$c_k = \frac{1}{\|\varphi_k\|^2} \int_a^b f(x)\varphi_k(x)\,dx,\quad \|\varphi_k\| = \left(\int_a^b \varphi_k^2(x)\,dx\right)^{1/2} \tag{11.136}$$

is convergent in the mean to $f(x)$ on $[a, b]$, that is

$$\rho^2\left(f, \sum_{k=1}^{n} c_k\varphi_k\right) = \int_a^b \left[f(x) - \sum_{k=1}^{n} c_k\varphi_k(x)\right]^2 dx \to 0 \quad \text{for } n \to +\infty \tag{11.137}$$

In this case we say that system (11.134) is a **basis of the functional** space $Q[a, b]$ since, in the case of completeness, for each "element" $f(x) \in Q[a, b]$ we can write the generalized equality

$$f(x) = \sum_{k=1}^{+\infty} c_k\varphi_k(x) \quad \text{where} \quad c_k = \frac{1}{\|\varphi_k\|^2}\int_a^b f(x)\varphi_k(x)\,dx \tag{11.138}$$

which should be understood in the sense of convergence in the mean on the interval $[a, b]$, i.e. in the sense that relation (11.137) is fulfilled (see § 6, Sec. 1 of Chapter 8).

2. Parseval Relation as a Necessary and Sufficient Condition for an Orthogonal System Being Complete. Let us make use of Bessel's identity

$$\rho^2\left(f, \sum_{k=1}^{n} c_k\varphi_k\right) = \int_a^b \left[f(x) - \sum_{k=1}^{n} c_k\varphi_k(x)\right]^2 dx =$$

$$= \int_a^b f^2(x)\,dx - \sum_{k=1}^{n} c_k^2 \|\varphi_k\|^2 \tag{11.139}$$

(see § 3, Sec. 3). Passing to the limit in (11.139) for $n \to +\infty$ we obtain

$$\lim_{n\to+\infty} \rho^2\left(f, \sum_{k=1}^{n} c_k\varphi_k\right) = \int_a^b f^2(x)\,dx - \sum_{k=1}^{+\infty} c_k^2\|\varphi_k\|^2 \tag{11.140}$$

whence it follows that the relation

$$\lim_{n\to+\infty} \rho^2\left(f, \sum_{k=1}^{n} c_k\varphi_k\right) = 0 \tag{11.141}$$

is equivalent to the equality

$$\int_a^b f^2(x)\,dx = \sum_{k=1}^{+\infty} c_k^2 \|\varphi_k\|^2 \tag{11.142}$$

Equality (11.142) is known as **Parseval's relation**. Thus, orthogonal system (11.134) is complete if and only if for any function $f(x) \in \in Q[a, b]$ Parseval's relation (11.142) is fulfilled.

3. Properties of Complete Systems. *An orthogonal system*

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots \tag{11.143}$$

*on an interval $[a, b]$ is said to be **closed** if every function $f(x) \in \in Q[a, b]$ which is orthogonal to all the functions of system (11.143) is the zero element of the functional space $Q[a, b]$, i.e. $f(x)$ is equal to zero at all its points of continuity and hence is different from zero only at a finite number of points of the interval $[a, b]$.*

Theorem 11.6. If system (11.143) is orthogonal and complete on $[a, b]$ it is closed.

Proof. Let a piecewise continuous function $f(x)$ be orthogonal on $[a, b]$ to all the functions of system (11.143), i.e.

$$\int_a^b f(x)\varphi_k(x)\,dx = 0 \quad \text{for } k = 1, 2, \ldots$$

Then the Fourier coefficients of the function $f(x)$ with respect to system (11.143) are equal to zero:

$$c_k = \frac{1}{\|\varphi_k\|^2}\int_a^b f(x)\varphi_k(x)\,dx = 0,\ \ k = 1, 2, \ldots \tag{11.144}$$

By the completeness of system (11.143), Parseval's relation

$$\int_a^b f^2(x)\,dx = \sum_{k=1}^{+\infty} c_k^2 \|\varphi_k\|^2 \tag{11.145}$$

is fulfilled for any function $f(x) \in Q[a, b]$. But then, by virtue of (11.144), equality (11.145) implies that

$$\int_a^b f^2(x)\,dx = 0 \tag{11.146}$$

Suppose that $f(x_0) \neq 0$ where $x_0 \in [a, b]$ is a point at which $f(x)$ is continuous. Let us embed the point x_0 in an interval $[a', b']$ (lying within $[a, b]$) on which $f(x)$ is continuous. The function $f^2(x)$ being continuous and nonnegative on $[a', b']$, we have $\int\limits_{a'}^{b'} f^2(x)\,dx > 0$ since $f^2(x_0) > 0$. But this implies that $\int\limits_a^b f^2(x)\,dx > 0$ which contradicts equality (11.146). Consequently, $f(x) \equiv 0$ at all the points of continuity on $[a, b]$, and thus the theorem has been proved.

Theorem 11.7. *If two functions $f(x)$ and $g(x)$ belonging to $Q[a, b]$ have the same Fourier series with respect to complete orthogonal system* (11.143) *on $[a, b]$, these functions coincide as elements of the space $Q[a, b]$, that is they may differ only at a finite number of points of the interval $[a, b]$.*

Proof. The function $\psi(x) = (f(x) - g(x)) \in Q[a, b]$ is orthogonal to all the functions of system (11.143) on $[a, b]$. In fact, we have

$$\int\limits_a^b \psi(x)\,\varphi_k(x)\,dx = \int\limits_a^b f(x)\,\varphi_k(x)\,dx - \int\limits_a^b g(x)\,\varphi_k(x)\,dx =$$

$$= c_k^f \,\|\varphi_k\|^2 - c_k^g\,\|\varphi_k\|^2 = (c_k^f - c_k^g)\,\|\varphi_k\|^2, \quad k = 1, 2, 3, \ldots \tag{11.147}$$

where c_k^f are the Fourier coefficients of the function $f(x)$ and c_k^g are the Fourier coefficients of the function $g(x)$. By the hypothesis, the Fourier series of these functions coincide, i.e. $c_k^f = c_k^g$ for $k = 1, 2, \ldots$, and hence it follows from (11.147) that

$$\int\limits_a^b \psi(x)\,\varphi_k(x)\,dx = 0 \quad \text{for } k = 1, 2, 3, \ldots \tag{11.148}$$

But then, by the foregoing theorem, the difference $\psi(x) = f(x) - g(x)$ is identically equal to zero on $[a, b]$ at all the points of continuity of $\psi(x)$ and thus may be different from zero only at a finite number of points of the interval $[a, b]$ which is what we set out to prove.

Theorem 11.8. *If system* (11.143) *defined on the interval $[a, b]$ is orthogonal and complete on $[a, b]$, then, for any two functions $f(x)$ and $g(x)$ belonging to $Q[a, b]$ we have the* ***generalized Parseval relation***

$$\int\limits_a^b f(x)\,g(x)\,dx = \sum_{k=1}^{+\infty} c_k^f c_k^g\,\|\varphi_k\|^2 \tag{11.149}$$

where c_k^f (c_k^g) are the Fourier coefficients of $f(x)$ ($g(x)$) with respect to orthogonal system (11.143).

Proof. Equality (11.149) is obtained if we write Parseval's relation for the functions $f(x)+g(x)$ and $f(x)-g(x)$ and then subtract the latter from the former and take half of the result.

Theorem 11.9. Let $f(x) \in Q[a, b]$. If an orthogonal system of functions $\{\varphi_i(x)\}$ defined on $[a, b]$ is complete, the Fourier series of the function $f(x)$ with respect to the system $\{\varphi_i(x)\}$ can be integrated term-by-term, that is

$$\int_{x_0}^{x} f(\xi)\,d\xi=\sum_{k=1}^{+\infty} c_k \int_{x_0}^{x} \varphi_k(\xi)\,d\xi \tag{11.150}$$

for any x_0 and x belonging to the interval $[a, b]$.

Proof. This theorem follows from the fact that the Fourier series $\sum\limits_{k=1}^{+\infty} c_k\varphi_k(x)$ converges in the mean to $f(x)$ on $[a, b]$. Indeed, as was proved, it is allowable to integrate termwise the series which converge in the mean (see Theorem 8.14_2 in Chapter 8, § 6, Sec. 3).

4. Completeness of Trigonometric System.

Theorem 11.10. The trigonometric system

$$\frac{1}{2},\ \cos\frac{\pi x}{l},\ \sin\frac{\pi x}{l},\ \ldots,\ \cos\frac{k\pi x}{l},\ \sin\frac{k\pi x}{l}\ \ldots \tag{11.151}$$

is complete.

Proof. It is required to prove that the relation

$$\rho^2(f, T_n^f)=\int_{-l}^{l}[f(x)-T_n^f(x)]^2\,dx \to 0 \quad \text{for } n\to+\infty \tag{11.152}$$

holds for every piecewise continuous function $f(x)$ defined on $[-l, l]$ where

$$T_n^f(x)=\frac{a_0}{2}+\sum_{k=1}^{n}\left(a_k\cos\frac{k\pi x}{l}+b_k\sin\frac{k\pi x}{l}\right)$$

is the Fourier polynomial of the function $f(x)$ with respect to system (11.151).

Let $|f(x)|<M$ on $[-l, l]$ and let $\varepsilon>0$ be an arbitrary number. Without loss of generality, we can suppose that $f(x)$ has a single point of discontinuity x_0 lying in the interior of $[-l, l]$. We shall construct a continuous function $g(x)$ on the interval $[-l, l]$ such

that it assumes equal values $g(-l) = g(l)$ at the end points of this interval and satisfies the inequality

$$\rho^2(f, g) = \int_{-l}^{l} [f(x) - g(x)]^2\, dx < \frac{\varepsilon}{4} \tag{11.153}$$

For this purpose we take a sufficiently small $\delta > 0$ and put $g(x) = f(x)$ for $-l \leqslant x \leqslant x_0 - \delta$ and for $x_0 + \delta \leqslant x \leqslant l - \delta$, $g(l) = f(-l)$, and consider $g(x)$ being linear on the intervals $x_0 - \delta \leqslant x \leqslant x_0 + \delta$ and $l - \delta \leqslant x \leqslant l$ (see Fig. 11.10 where the graph

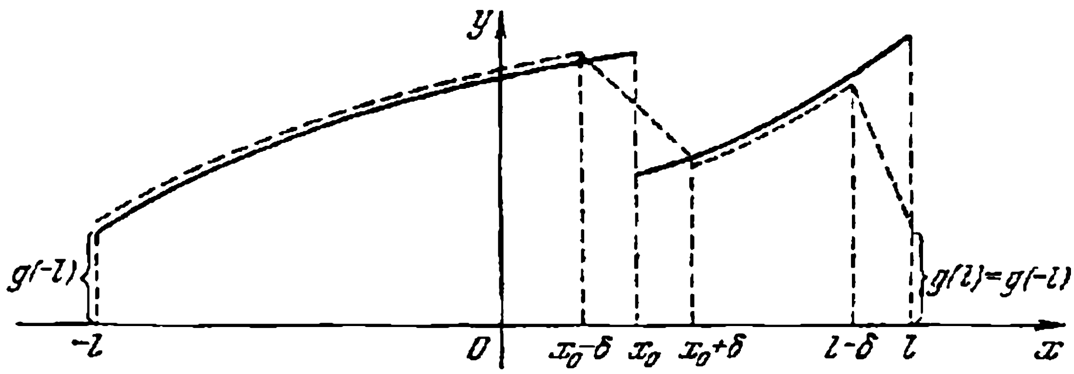


Fig. 11.10

of $f(x)$ is shown in the continuous line and the graph of $g(x)$ in the dotted line). According to the way $g(x)$ has been constructed, we have $g(-l) = g(l) = f(-l)$, and the difference $f(x) - g(x)$ may be different from zero only for $x_0 - \delta < x < x_0 + \delta$ and $l - \delta < x < l$. Therefore we can write

$$\begin{aligned}\rho^2(f, g) &= \int_{-l}^{l} [f(x) - g(x)]^2\, dx = \int_{x_0-\delta}^{x_0+\delta} [f(x) - g(x)]^2\, dx \\ &+ \int_{l-\delta}^{l} [f(x) - g(x)]^2\, dx \leqslant \int_{x_0-\delta}^{x_0+\delta} \{|f(x)| + |g(x)|\}^2\, dx + \\ &+ \int_{l-\delta}^{l} \{|f(x)| + |g(x)|\}^2\, dx \leqslant 4M^2 2\delta + 4M^2\delta = 12M^2\delta < \frac{\varepsilon}{4}\end{aligned}$$

provided that $\delta > 0$ is sufficiently small. Since the function $g(x)$ is continuous on the interval $[-l, l]$ and takes equal values at its end points ($g(-l) = g(l)$), Theorem 11.4 (Weierstrass' trigonometric approximation theorem) implies that there is a trigonometric polynomial

$$T_{n_0}(x) = \frac{\alpha_0}{2} + \sum_{k=1}^{n} \left(\alpha_k \cos\frac{k\pi x}{l} + \beta_k \sin\frac{k\pi x}{l}\right)$$

such that

$$|g(x) - T_{n_0}(x)| < \sqrt{\frac{\varepsilon}{8l}} \text{ for all } x \in [-l, l] \tag{11.154}$$

Consequently,

$$\rho^2(g, T_{n_0}) = \int_{-l}^{l} |g(x) - T_{n_0}(x)|^2\,dx < \frac{\varepsilon}{8l}\int_{-l}^{l} dx = \frac{\varepsilon}{4} \qquad (11.155)$$

Now taking advantage of the inequality

$$(a+b)^2 \leqslant 2a^2 + 2b^2$$

and putting $a = f(x) - g(x)$ and $b = g(x) - T_{n_0}(x)$ we obtain

$$\{f(x) - T_{n_0}(x)\}^2 \leqslant 2\,\{[f(x) - g(x)]^2 + [g(x) - T_{n_0}(x)]^2\}$$

It follows that

$$\rho^2(f, T_{n_0}) = \int_{-l}^{l} [f(x) - T_{n_0}(x)]^2\,dx \leqslant 2\int_{-l}^{l} [f(x) - g(x)]^2\,dx$$

$$+\,2\int_{-l}^{l} [g(x) - T_{n_0}(x)]^2\,dx < 2\,\frac{\varepsilon}{4} + 2\,\frac{\varepsilon}{4} = \varepsilon$$

If we substitute the Fourier trigonometric polynomial $T'_{n_0}(x)$ corresponding to the function $f(x)$ for the trigonometric polynomial $T_{n_0}(x)$ in the last inequality we shall have

$$\rho^2(f, T'_{n_0}) \leqslant \varepsilon \qquad (11.156)$$

because T'_{n_0} gives the minimum to the mean square deviation.

Making use of Bessel's identity we can rewrite inequality (11.156) in the form

$$\rho^2(f, T'_{n_0}) = \int_{-l}^{l} f^2(x)\,dx - l\left\{\frac{a_0^2}{2} + \sum_{k=1}^{n_0}(a_k^2 + b_k^2)\right\} < \varepsilon \qquad (11.157)$$

Consequently, we have

$$\rho^2(f, T'_n) = \int_{-l}^{l} f^2(x)\,dx - l\left\{\frac{a_0^2}{2} + \sum_{k=1}^{n}(a_k^2 + b_k^2)\right\} < \varepsilon \qquad (11.158)$$

for all $n \geqslant n_0$. Since $\varepsilon > 0$ has been chosen quite arbitrarily, it follows that $\rho^2(f, T'_n) \to 0$ for $n \to +\infty$ which is what we set out to prove.

The completeness of the trigonometric system which has been established here implies that this system is closed. It also means that a piecewise continuous function $f(x)$ is uniquely specified by its Fourier series with respect to the trigonometric system everywhere on the interval $[-l, l]$ except possibly at a finite number of points (at which $f(x)$ is discontinuous). The completeness of the

trigonometric system was for the first time proved by A. M. Lyapunov*.

5. Completeness of Some Other Classical Orthogonal Systems. In mathematical physics we deal with various orthogonal systems, other than the trigonometric system, whose completeness is established similarly. As an example, let us prove that the system of Legendre's polynomials is complete. Consider a piecewise continuous function $f(x)$ defined on the interval $[-1, 1]$. Suppose we are given an arbitrary $\varepsilon > 0$. By analogy with the proof of the completeness of the trigonometric system, let us construct a function $g_\varepsilon(x)$ which is continuous on $[-1, 1]$ and satisfies the relation

$$\rho^2(f, g_\varepsilon) = \int_{-1}^{1} [f(x) - g_\varepsilon(x)]^2\, dx < \frac{\varepsilon}{4} \tag{11.159}$$

(but here the requirement that $g_\varepsilon(x)$ must assume equal values at the end points of the interval $[-1, 1]$ is dropped). Theorem 11.5 (Weierstrass' polynomial approximation theorem) indicates that there exists an algebraic polynomial

$$Q_m(x) = A_0 + A_1 x + A_2 x^2 + \ldots + A_m x^m$$

satisfying the inequality

$$|g_\varepsilon(x) - Q_m(x)| < \frac{1}{2}\sqrt{\frac{\varepsilon}{2}}$$

for all $-1 \leqslant x \leqslant 1$. This implies the inequality

$$\rho^2(g_\varepsilon, Q_m) = \int_{-1}^{1} [g_\varepsilon(x) - Q_m(x)]^2\, dx < \frac{\varepsilon}{4} \tag{11.160}$$

The functions $1, x, x^2, \ldots, x^m$ are linear combinations of Legendre's polynomials** and therefore

$$Q_m(x) = B_0 + B_1 P_1(x) + B_2 P_2(x) + \ldots + B_m P_m(x)$$

where $P_1(x), \ldots, P_m(x)$ are Legendre's polynomials.

Taking advantage of the inequality $(a + b)^2 \leqslant 2a^2 + 2b^2$ we conclude, on the basis of relations (11.159) and (11.160), that

$$\rho^2(f, Q_m) \leqslant 2\rho^2(f, g_\varepsilon) + 2\rho^2(g_\varepsilon, Q_m) \leqslant 2\frac{\varepsilon}{4} + 2\frac{\varepsilon}{4} = \varepsilon \tag{11.161}$$

Let us substitute the Fourier coefficients of the function $f(x)$ with respect to the system of Legendre's polynomials for the coefficients

* Lyapunov, Aleksandr Mikhailovich (1857-1918), a prominent Russian mathematician.

** See Appendix 1 to Chapter 11.

$B_0, \ldots, B_m$ in the expression of $\rho^2(f, Q_m)$, i.e. replace B_k, $k = 0, 1, \ldots, m$, by the quantities

$$c_k = \frac{1}{\|P_k\|^2} \int_{-1}^{1} f(x) P_k(x)\, dx, \quad k = 0, 1, \ldots, m \qquad (11.162)$$

Since Fourier's polynomials minimize the mean square deviation, the above substitution cannot increase the mean square deviation. Therefore, introducing the notation

$$Q_m^f(x) = c_0 + c_1 P_1(x) + \ldots + c_m P_m(x) \qquad (11.163)$$

we arrive at the inequality

$$\rho^2(f, Q_m^f) < \varepsilon \qquad (11.164)$$

Furthermore, by Bessel's identity

$$\rho^2(f, Q_n^f) = \int_{-1}^{1} f^2(x)\, dx - \sum_{k=1}^{n} c_k^2 \|\varphi_k\|^2 \qquad (11.165)$$

we conclude that relation (11.164) implies the validity of the inequality

$$\rho^2(f, Q_n^f) < \varepsilon \qquad (11.166)$$

for all $n \geqslant m$. Since $\varepsilon > 0$ has been taken arbitrarily, it follows that

$$\rho^2(f, Q_n^f) \to 0 \quad \text{for} \quad n \to +\infty \qquad (11.167)$$

and thus we have established the completeness of the system of Legendre's polynomials.

§ 7. FOURIER SERIES IN ORTHOGONAL SYSTEMS OF COMPLEX FUNCTIONS

Here, besides real functions, we shall also consider complex functions of a real independent variable x which are of the form

$$\varphi(x) = \varphi^*(x) + i\varphi^{**}(x) \qquad (11.168)$$

where $\varphi^*(x)$ and $\varphi^{**}(x)$ are real functions. The function which is the complex conjugate of $\varphi(x)$ (i.e. the one which differs from $\varphi(x)$ only in the sign of its imaginary part) will be denoted by $\overline{\varphi}(x)$. Thus,

$$\overline{\varphi}(x) = \varphi^*(x) - i\varphi^{**}(x) \qquad (11.168')$$

It should be noted that

$$\varphi(x)\overline{\varphi}(x) = [\varphi^*(x)]^2 + [\varphi^{**}(x)]^2 = |\varphi(x)|^2 \geqslant 0 \qquad (11.169)$$

A function $\varphi(x) = \varphi^*(x) + i\varphi^{**}(x)$ is said to be *continuous* (*piecewise continuous*) on $[a, b]$ if its real and imaginary parts, that

is the functions $\varphi^*(x)$ and $\varphi^{**}(x)$, are continuous (piecewise continuous) on $[a, b]$.

The derivative and the integral of a function $\varphi(x) = \varphi^*(x) + i\varphi^{**}(x)$ are defined, respectively, by the equalities

$$\frac{d\varphi}{dx} = \frac{d\varphi^*}{dx} + i\,\frac{d\varphi^{**}}{dx} \tag{11.170}$$

and

$$\int_a^b \varphi(x)\,dx = \int_a^b \varphi^*(x)\,dx + i\int_a^b \varphi^{**}(x)\,dx \tag{11.171}$$

and the function $\varphi(x) = \varphi^*(x) + i\varphi^{**}(x)$ is said to be *differentiable* (*integrable*) if $\varphi^*(x)$ and $\varphi^{**}(x)$ are differentiable (integrable).

If two functions $\varphi(x) = \varphi^*(x) + i\varphi^{**}(x)$ and $\psi(x) = \psi^*(x) + i\psi^{**}(x)$ are integrable on $[a, b]$, it is obvious that the function $\varphi(x)\overline{\psi}(x)$ is also integrable on $[a, b]$. In particular, if $\varphi(x)$ is integrable on $[a, b]$ the function $\varphi(x)\overline{\varphi}(x)$ is also integrable on $[a, b]$ and

$$\int_a^b \varphi(x)\,\overline{\varphi}(x)\,dx = \int_a^b |\varphi(x)|^2\,dx = \int_a^b \{[\varphi^*(x)]^2 + [\varphi^{**}(x)]^2\}\,dx \geqslant 0$$

Two functions $\varphi(x)$ and $\psi(x)$ integrable on an interval $[a, b]$ are said to be **orthogonal** on this interval if

$$\int_a^b \varphi(x)\,\overline{\psi}(x)\,dx = 0 \tag{11.172}$$

A system of complex functions

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots \tag{11.173}$$

integrable on $[a, b]$ is called **orthogonal** on $[a, b]$ if

$$\int_a^b \varphi_j(x)\,\overline{\varphi}_k(x)\,dx = \begin{cases} 0 & \text{for } j \neq k \\ \|\varphi_k\|^2 > 0 & \text{for } j = k \end{cases} \tag{11.174}$$

where $\|\varphi_k\| = \int_a^b |\varphi_k(x)|^2\,dx$ is the *norm* of φ_k. Generally, the norm of an integrable complex function $\varphi(x)$ is defined as the nonnegative quantity

$$\|\varphi\| = \left(\int_a^b \varphi(x)\,\overline{\varphi}(x)\,dx\right)^{1/2} = \left(\int_a^b |\varphi(x)|^2\,dx\right)^{1/2} \tag{11.175}$$

One of the most important examples of orthogonal systems of complex functions is the system

$$e^{i\frac{n\pi x}{l}}, \quad n=0, \pm 1, \pm 2, \ldots \tag{11.176}$$

which is orthogonal on the interval $[-l, l]$. The orthogonality of $e^{i\frac{k\pi x}{l}}$ and $e^{i\frac{n\pi x}{l}}$ for $k \neq n$ is established directly by integrating the product

$$e^{i\frac{k\pi x}{l}}\left(\overline{e^{i\frac{n\pi x}{l}}}\right)=e^{i\frac{k\pi x}{l}}e^{-i\frac{n\pi x}{l}}=e^{i\frac{(k-n)\pi x}{l}}=\cos\frac{(k-n)\pi x}{l}+$$
$$-i\sin\frac{(k-n)\pi x}{l}$$

over the interval $[-l, l]$. For the norm of the function $e^{i\frac{n\pi x}{l}}$ we obtain the expression

$$\left\|e^{i\frac{n\pi x}{l}}\right\|=\left(\int_{-l}^{l}e^{i\frac{n\pi x}{l}}e^{-i\frac{n\pi x}{l}}dx\right)^{1/2}=\left(\int_{-l}^{l}dx\right)^{1/2}=\sqrt{2l} \tag{11.177}$$

The *Fourier coefficients* of a function $f(x)$ (integrable on $[a, b]$) with respect to orthogonal system (11.173) are determined by the formulas

$$c_k=\frac{1}{\|\varphi_k\|^2}\int_a^b f(x)\,\overline{\varphi_k}(x)\,dx, \qquad k=1, 2, \ldots \tag{11.178}$$

The *Fourier series* of an integrable function $f(x)$ with respect to orthogonal system (11.173) is, by definition, the series

$$f(x)\sim\sum_{k=1}^{+\infty}c_k\varphi_k(x) \tag{11.179}$$

whose coefficients c_k are determined by formulas (11.178).

In particular, the Fourier coefficients of $f(x)$ with respect to system (11.176) are equal to

$$c_k=\frac{1}{2l}\int_{-l}^{l}f(x)\,e^{-i\frac{k\pi x}{l}}dx, \quad k=0, \pm 1, \pm 2, \ldots \tag{11.178'}$$

and the corresponding Fourier series of $f(x)$ with respect to this system is the two-way series

$$f(x)\sim\sum_{n=-\infty}^{+\infty}c_n e^{i\frac{n\pi x}{l}} \tag{11.179'}$$

Let us prove that if the function $f(x)$ is real on the interval $[-l, l]$, relations (11.178′) and (11.179′) are equivalent to the relations

$$\begin{aligned} a_k &= \frac{1}{l}\int_{-l}^{l} f(x)\cos\frac{k\pi x}{l}\,dx, \quad k=0,\ 1,\ 2,\ \ldots \\ b_k &= \frac{1}{l}\int_{-l}^{l} f(x)\sin\frac{k\pi x}{l}\,dx, \quad k=1,\ 2,\ \ldots \end{aligned} \tag{11.180}$$

and

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{+\infty}\left(a_k\cos\frac{k\pi x}{l} + b_k\sin\frac{k\pi x}{l}\right) \tag{11.181}$$

that is relations (11.178′) and (11.179′) are the complex form of Fourier's coefficients and series of the function $f(x)$ with respect to the trigonometric system.

Applying Euler's formula

$$e^{i\varphi} = \cos\varphi + i\sin\varphi$$

to (11.178′) we obtain

$$c_0 = \frac{1}{2l}\int_{-l}^{l} f(x)\,dx = \frac{a_0}{2} \tag{11.182}$$

$$\begin{aligned} c_k &= \frac{1}{2l}\int_{-l}^{l} f(x)\,e^{-i\frac{k\pi x}{l}}\,dx = \\ &= \frac{1}{2l}\int_{-l}^{l} f(x)\left[\cos\frac{k\pi x}{l} - i\sin\frac{k\pi x}{l}\right]dx = \\ &= \frac{a_k - ib_k}{2}, \quad k=1,\ 2,\ \ldots \end{aligned} \tag{11.183}$$

an

$$\begin{aligned} c_{-k} &= \frac{1}{2l}\int_{-l}^{l} f(x)\,e^{i\frac{k\pi x}{l}}\,dx = \frac{1}{2l}\int_{-l}^{l} f(x)\left[\cos\frac{k\pi x}{l} + \right. \\ &\left. + i\sin\frac{k\pi x}{l}\right]dx = \frac{a_k + ib_k}{2}, \quad k=1,\ 2,\ \ldots \end{aligned} \tag{11.184}$$

Series (11.179′) can be rewritten as

$$\sum_{n=-\infty}^{+\infty} c_n e^{i\frac{n\pi x}{l}} = c_0 + \sum_{k=1}^{+\infty} c_k e^{i\frac{k\pi x}{l}} + \sum_{k=1}^{+\infty} c_{-k} e^{-i\frac{k\pi x}{l}} \tag{11.185}$$

Substituting expressions (11.182), (11.183) and (11.184) of the coefficients c_0, c_k and c_{-k} into (11.185) and taking advantage of

Euler's formulas

$$\cos\varphi=\frac{e^{i\varphi}+e^{-i\varphi}}{2},\quad \sin\varphi=\frac{e^{i\varphi}-e^{-i\varphi}}{2i}$$

we arrive at the equality

$$\sum_{n=-\infty}^{+\infty}{}' c_n e^{i\frac{n\pi x}{l}}=\frac{a_0}{2}+\sum_{k=1}^{+\infty}\frac{a_k-ib_k}{2}e^{i\frac{k\pi x}{l}}+\sum_{k=1}^{+\infty}\frac{a_k-ib_k}{2}e^{-i\frac{k\pi x}{l}}=$$

$$=\frac{a_0}{2}+\sum_{k=1}^{+\infty}\left[a_k\frac{e^{i\frac{k\pi x}{l}}+e^{-i\frac{k\pi x}{l}}}{2}+\right.$$

$$\left.+b_k\frac{e^{i\frac{k\pi x}{l}}-e^{-i\frac{k\pi x}{l}}}{2i}\right]=\frac{a_0}{2}+$$

$$+\sum_{k=1}^{+\infty}\left(a_k\cos\frac{k\pi x}{l}+b_k\sin\frac{k\pi x}{l}\right) \qquad (11.186)$$

which is what we set out to prove.

If the function $f(x)$ is not only integrable but also piecewise smooth on $[-l, l]$, the theorem on convergence of a Fourier series and relation (11.186) enable us to write down the expansion

$$f(x)=\sum_{n=-\infty}^{+\infty}{}' c_n e^{i\frac{n\pi x}{l}} \qquad (11.187)$$

where the Fourier coefficients are determined by the formulas (11.178). At the points of discontinuity of the function $f(x)$ the left-hand side of equality (11.187) should be replaced by $\frac{f(x+0)+f(x-0)}{2}$ if $-l<x<l$ and by $\frac{f(l-0)+f(-l+0)}{2}$ if $x=\pm l$.

Complex form (11.178′), (11.179′) of expansion of functions into trigonometric series is widely used in mathematical physics and its applications. This form is especially convenient for performing calculations and transformations, in particular, when the expressions we deal with involve the products of Fourier's series and also for Fourier's series of functions of several independent variables.

§ 8. FOURIER SERIES FOR FUNCTIONS OF SEVERAL INDEPENDENT VARIABLES

Let a function $f(x, y)$ be defined in a rectangle $-l_1\leqslant x\leqslant l_1$, $-l_2\leqslant y\leqslant l_2$. We shall suppose that for every $y\in[-l_2, l_2]$ the function $f(x, y)$ satisfies the conditions under which it can be expan-

ded into Fourier's series as a function of x on the interval $[-l_1, l_2]$. Then, using the complex form of Fourier's series we can write

$$f(x, y)=\sum_{n=-\infty}^{+\infty} c_n(y)\, e^{i\frac{n\pi x}{l_1}} \tag{11.188}$$

where

$$c_n(y)=\frac{1}{2l_1}\int_{-l_1}^{l_1} f(\xi, y)\, e^{-i\frac{n\pi\xi}{l_1}}\, d\xi, \quad n=0, \pm 1, \pm 2, \ldots \tag{11.189}$$

Let each of the functions $c_n(y)$, in its turn, have the Fourier series expansion

$$c_n(y)=\sum_{m=-\infty}^{+\infty} c_{nm} e^{i\frac{m\pi y}{l_2}}, \quad n=0, \pm 1, \pm 2, \ldots \tag{11.190}$$

on the interval $-l_2 \leqslant y \leqslant l_2$. The coefficients c_{nm} are determined by the formulas

$$c_{nm}=\frac{1}{2l_2}\int_{-l_2}^{l_2} c_n(\eta)\, e^{-i\frac{m\pi\eta}{l_2}}\, d\eta, \quad n, m=0, \pm 1, \pm 2, \ldots \tag{11.191}$$

Now, substituting (11.189) into (11.191) and (11.190) into (11.188) we obtain

$$f(x, y)=\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} c_{nm} e^{i\left(\frac{n\pi x}{l_1}+\frac{m\pi y}{l_2}\right)} \tag{11.192}$$

where

$$c_{nm}=\frac{1}{4l_1l_2}\int_{-l_1}^{l_1}\int_{-l_2}^{l_2} f(\xi, \eta)\, e^{-i\left(\frac{n\pi\xi}{l_1}+\frac{m\pi\eta}{l_2}\right)}\, d\xi\, d\eta \tag{11.193}$$

Thus, we have obtained, in complex form, the Fourier series expansion for the function $f(x, y)$ of two independent variables

Making use of Euler's formula $e^{i\varphi}=\cos\varphi+i\sin\varphi$ we can rewrite expansion (11.192) in the form

$$\begin{aligned} f(x, y) &= \sum_{n, m=-\infty}^{+\infty} c_{nm}\left(\cos\frac{n\pi x}{l_1}+i\sin\frac{n\pi x}{l_1}\right)\left(\cos\frac{m\pi y}{l_2}+\right. \\ &\left.+ i\sin\frac{m\pi y}{l_2}\right)=\sum_{m, n=-\infty}^{+\infty}\lambda_{mn}\left[a_{mn}\cos\frac{n\pi x}{l_1}\cos\frac{m\pi y}{l_2}+\right. \\ &+ b_{mn}\sin\frac{n\pi x}{l_1}\cos\frac{m\pi y}{l_2}+c'_{mn}\cos\frac{n\pi x}{l_1}\sin\frac{m\pi y}{l_2}+ \\ &\left.+ d_{mn}\sin\frac{n\pi x}{l_1}\sin\frac{m\pi y}{l_2}\right] \end{aligned} \tag{11.194}$$

where

$$\lambda_{mn} = \begin{cases} \frac{1}{4} & \text{for } m = n = 0 \\ \frac{1}{2} & \text{for } m = 0,\ n > 0 \text{ and } m > 0,\ n = 0 \\ 1 & \text{for } m > 0,\ n > 0 \end{cases} \tag{11.195}$$

and

$$\begin{aligned} a_{mn} &= \frac{1}{l_1 l_2} \int_{-l_1}^{l_1} \int_{-l_2}^{l_2} f(x,\ y) \cos \frac{n\pi x}{l_1} \cos \frac{m\pi y}{l_2}\, dx\, dy \\ b_{mn} &= \frac{1}{l_1 l_2} \int_{-l_1}^{l_1} \int_{-l_2}^{l_2} f(x,\ y) \sin \frac{n\pi x}{l_1} \cos \frac{m\pi y}{l_2}\, dx\, dy \\ c_{mn} &= \frac{1}{l_1 l_2} \int_{-l_1}^{l_1} \int_{-l_2}^{l_2} f(x,\ y) \cos \frac{n\pi x}{l_1} \sin \frac{m\pi y}{l_2}\, dx\, dy \\ d_{mn} &= \frac{1}{l_1 l_2} \int_{-l_1}^{l_1} \int_{-l_2}^{l_2} f(x,\ y) \sin \frac{n\pi x}{l_1} \sin \frac{m\pi y}{l_2}\, dx\, dy \end{aligned} \tag{11.196}$$

If $f(x,\ y)$ is an even function with respect to each argument, that is if

$$f(-x,\ y) \equiv f(x,\ -y) = f(x,\ y) \tag{11.197}$$

we can easily see that $b_{mn} = c_{mn} = d_{mn} = 0$ and therefore the Fourier series of such a function takes the form

$$f(x,\ y) = \sum_{m,\ n=0}^{+\infty} \lambda_{mn} a_{mn} \cos \frac{n\pi x}{l_1} \cos \frac{m\pi y}{l_2} \tag{11.198}$$

If $f(x,\ y)$ is odd with respect to each argument x and y, only its coefficients d_{mn} can be different from zero and, consequently, its Fourier series is of the form

$$f(x,\ y) = \sum_{m,\ n=0}^{+\infty} d_{mn} \sin \frac{n\pi x}{l_1} \sin \frac{m\pi y}{l_2} \tag{11.199}$$

If $f(x,\ y)$ is even in y and odd in x its expansion contains only the functions $\sin \frac{n\pi x}{l_1} \cos \frac{m\pi y}{l_2}$ and if it is an odd function of y and an even function of x the expansion involves the functions $\cos \frac{n\pi x}{l_1} \sin \frac{m\pi y}{l_2}$.

Here we do not investigate the conditions guaranteeing the possibility of expanding a function of two variables $f(x, y)$ into a Fourier multiple series and limit ourselves to the formulation (without proof) of the following assertion: *if $f(x, y)$, $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial^2 f}{\partial x\,\partial y}$ are continuous functions periodic in x with period $2l_1$ and in y with period $2l_2$, the Fourier double series of the function $f(x, y)$ converges to $f(x, y)$ at every point.*

§ 9. FOURIER INTEGRAL

1. Formal Derivation of Fourier Integral Formula. In this section we present a formal derivation of *Fourier's integral formula* which is obtained if we take the Fourier series of a function $f(x)$ on an interval $[-l, l]$ and then extend this interval to infinity, i.e. pass to the limit as $l \to +\infty$. Then the Fourier series turns into the Fourier integral.

When we consider a function $f(x)$ defined on an arbitrary finite interval $[-l, l]$, it is expanded into a sum of "harmonic oscillations" whose frequencies form a descrete number sequence. But when we take the limit for $l \to +\infty$ we pass to a function defined over the whole x-axis (or on the semi-infinite interval $[0, +\infty]$ of the x-axis) which is presented by an integral, i.e. a sum of "harmonic vibrations" whose frequencies λ constitute the continuous interval $0 \leqslant \lambda < +\infty$. Let us consider this formal passage to the limit from the Fourier series to the Fourier integral.

Let $f(x)$ be defined on the entire x axis. We shall suppose that the function $f(x)$ is piecewise smooth on every finite interval $[-l, l]$. Then, by the convergence theorem for trigonometric series, we have, for every $l > 0$, the expansion

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{+\infty} \left(a_k \cos \frac{k\pi x}{l} + b_k \sin \frac{k\pi x}{l} \right) \tag{11.200}$$

where

$$\left.\begin{aligned} a_0 &= \frac{1}{l} \int_{-l}^{l} f(\xi)\, d\xi, \quad a_k = \frac{1}{l} \int_{-l}^{l} f(\xi) \cos \frac{k\pi\xi}{l}\, d\xi \\ b_k &= \frac{1}{l} \int_{-l}^{l} f(\xi) \sin \frac{k\pi\xi}{l}\, d\xi, \quad k = 1, 2, \ldots \end{aligned}\right\} \tag{11.201}$$

Equality (11.200) is valid if x is an interior point of the interval $[-l, l]$ at which $f(x)$ is continuous. If x is an interior point of the interval at which the function $f(x)$ has a discontinuity, the left-hand side of equality (11.200) must be replaced by the expression

$\frac{f(x+0)+f(x-0)}{2}$. Substituting formulas (11.201) into (11.200) we obtain

$$f(x)=\frac{1}{2l}\int_{-l}^{l} f(\xi)\,d\xi+\frac{1}{l}\sum_{k=1}^{+\infty}\int_{-l}^{l} f(\xi)\cos\frac{k\pi}{l}(\xi-x)\,d\xi \qquad (11.202)$$

If the function $f(x)$ is absolutely integrable over the whole x-axis, that is

$$\int_{-\infty}^{+\infty}|f(x)|\,dx=Q<+\infty \qquad (11.203)$$

then, by virtue of condition (11.203), the first summand on the right-hand side of (11.202) tends to zero for $l\to+\infty$. Consequently, we have

$$f(x)=\lim_{l\to+\infty}\frac{1}{l}\sum_{k=1}^{+\infty}\int_{-l}^{l} f(\xi)\cos\frac{k\pi}{l}(\xi-x)\,d\xi \qquad (11.204)$$

Putting $\frac{k\pi}{l}=\lambda_k$ and $\frac{\pi}{l}=\Delta\lambda_k$ we can rewrite relation (11.204) in the form

$$f(x)=\lim_{\substack{l\to+\infty\\ \Delta\lambda_k\to 0}}\frac{1}{\pi}\sum_{k=1}^{+\infty}\Delta\lambda_k\int_{-l}^{l} f(\xi)\cos\lambda_k(\xi-x)\,d\xi \qquad (11.205)$$

We can now conclude, intuitively (without any rigorous argument), that

(1) the integral $\int_{-l}^{l} f(\xi)\cos\lambda_k(\xi-x)\,d\xi$ can be replaced, for large values of l, by the integral $\int_{-\infty}^{+\infty} f(\xi)\cos\lambda_k(\xi-x)\,d\xi$, and

(2) the expression $\sum_{k=1}^{+\infty}\Delta\lambda_k\int_{-\infty}^{+\infty} f(\xi)\cos\lambda_k(\xi-x)\,d\xi$ is an integral sum corresponding to the integral $\int_{0}^{+\infty}d\lambda\int_{-\infty}^{+\infty} f(\xi)\cos\lambda(\xi-x)\,d\xi$, and thus, relation (11.205) implies the formula

$$f(x)=\frac{1}{\pi}\int_{0}^{+\infty}d\lambda\int_{-\infty}^{+\infty} f(\xi)\cos\lambda(\xi-x)\,d\xi \qquad (11.206)$$

If x is a point of discontinuity of $f(x)$, the left-hand side of (11.206) should be replaced by the expression $\frac{f(x+0)+f(x-0)}{2}$. Relati-

on (11.206) is called **Fourier's integral formula (theorem)** and the integral on the right-hand side of (11.206) is called the **Fourier integral.**

2. **Proof of Fourier Integral Theorem.** The above derivation of formula (11.206) is formal because it involves some passages to the limit which have not been justified rigorously. But it turns out that it is easier to present a direct proof of formula (11.206) than to justify these passages to the limit.

Theorem 11.11. If $f(x)$ is a piecewise smooth function on every finite interval of the x-axis absolutely integrable on the whole x-axis, i.e. $\int\limits_{-\infty}^{+\infty} |f(x)|\,dx$ *is convergent, then*

$$\lim_{l\to+\infty} \frac{1}{\pi}\int\limits_0^l d\lambda \int\limits_{-\infty}^{+\infty} f(\xi)\cos\lambda(\xi-x)\,d\xi = \frac{f(x+0)+f(x-0)}{2} \qquad (11.207)$$

Proof. Note that the integral $\int\limits_{-\infty}^{+\infty} f(\xi)\cos\lambda(\xi-x)\,d\xi$ dependent on the parameter λ is uniformly convergent with respect to λ for $0\leqslant\lambda<+\infty$ since $|f(\xi)\cos\lambda(\xi-x)|\leqslant|f(\xi)|$ and the integral $\int\limits_{-\infty}^{+\infty}|f(\xi)|\,d\xi$ is supposed to be convergent. Consequently, it is permissible to reverse the order of integration (see Sec. 3, § 2, Chapter 10) and rewrite the integral $\frac{1}{\pi}\int\limits_0^l d\lambda\int\limits_{-\infty}^{+\infty} f(\xi)\cos\lambda(\xi-x)\,d\xi$ in the form

$$\frac{1}{\pi}\int\limits_0^l d\lambda\int\limits_{-\infty}^{+\infty} f(\xi)\cos\lambda(\xi-x)\,d\xi = \frac{1}{\pi}\int\limits_{-\infty}^{+\infty} d\xi\int\limits_0^l f(\xi)\cos\lambda(\xi-x)\,d\lambda =$$
$$= \frac{1}{\pi}\int\limits_{-\infty}^{+\infty} f(\xi)\,\frac{\sin l(\xi-x)}{\xi-x}\,d\xi = \frac{1}{\pi}\int\limits_{-\infty}^{+\infty} f(x+\zeta)\,\frac{\sin l\zeta}{\zeta}\,d\zeta \qquad (11.208)$$

where $\zeta=\xi-x$, $d\zeta=d\xi$. Hence we must only prove that

$$\lim_{l\to+\infty}\frac{1}{\pi}\int\limits_{-\infty}^{0} f(x+\zeta)\,\frac{\sin l\zeta}{\zeta}\,d\zeta = \frac{f(x-0)}{2} \qquad (11.209)$$

and

$$\lim_{l\to+\infty}\frac{1}{\pi}\int\limits_{0}^{+\infty} f(x+\zeta)\,\frac{\sin l\zeta}{\zeta}\,d\zeta = \frac{f(x+0)}{2} \qquad (11.210)$$

For this purpose we shall take advantage of the relation

$$\frac{1}{\pi}\int_{0}^{+\infty}\frac{\sin l\zeta}{\zeta}\,d\zeta=\frac{1}{2} \qquad (11.21$$

which is implied by formula (10.50) derived in § 2, Sec. 5 Chapter 10.

As an instance, we shall establish the validity of relation (11.21 (formula (11.209) is proved similarly). By equality (11.211). can write

$$\frac{f(x+0)}{2}=\frac{1}{\pi}\int_{0}^{+\infty}f(x+0)\frac{\sin l\zeta}{\zeta}\,d\zeta \qquad (11.21$$

Therefore the difference between the variable $\frac{1}{\pi}\int_{0}^{\infty}f(x+\zeta)\frac{\sin l\zeta}{\zeta}$ and the constant $\frac{f(x+0)}{2}$ entering into relation (11.210) is equ to the expression

$$J_{0,+\infty}=\frac{1}{\pi}\int_{0}^{+\infty}f(x+\zeta)\frac{\sin l\zeta}{\zeta}\,d\zeta-\frac{f(x+0)}{2}=$$

$$=\frac{1}{\pi}\int_{0}^{+\infty}[f(x+\zeta)-f(x+0)]\frac{\sin l\zeta}{\zeta}\,d\zeta \qquad (11.21$$

Thus, we must show that the integral on the right-hand side (11.213) tends to zero as $l\to+\infty$. For this purpose we divide th interval of integration $0\leqslant\zeta<+\infty$ into three parts, namel $0\leqslant\zeta\leqslant\delta$, $\delta\leqslant\zeta\leqslant\Delta$ and $\Delta\leqslant\zeta<+\infty$. Then the integr we are interested in is presented as the sum of the correspondi integrals:

$$J_{0,+\infty}=J_{0,\delta}+J_{\delta,\Delta}+J_{\Delta,+\infty} \qquad (11.21$$

We shall first take an arbitrary $\varepsilon>0$ and prove that for all su ficiently small $\delta>0$ and all sufficiently large $\Delta>\delta$ we have th inequalities

$$|J_{0,\delta}|<\frac{\varepsilon}{3}\ \text{and}\ |J_{\Delta,+\infty}|<\frac{\varepsilon}{3} \qquad (11.21$$

which are fulfilled for all $l\geqslant1$ simultaneously. Then we shall fi some values of δ and Δ for which inequalities (11.215) hold an then choose a sufficiently large $l\geqslant1$ for which the relation $|J_{\delta,\Delta}|<$ $<\frac{\varepsilon}{3}$ is satisfied. The key lemma (see § 5) indicates that suc a choice of $l\geqslant1$ is possible. Then, by (11.214), we shall conclud that $|J_{0,+\infty}|<\varepsilon$ for all sufficiently large $l\geqslant1$

Thus, let us begin with estimating the integral

$$J_{0,\delta}=\frac{1}{\pi}\int_0^{\delta}\frac{f(x+\zeta)-f(x+0)}{\zeta}\sin l\zeta\,d\zeta$$

For all sufficiently small $\delta>0$ we have

$$\left|\frac{f(x+\zeta)-f(x+0)}{\zeta}\right|<|f_R'(x)|+1 \text{ for all } \zeta\in(0,\delta)$$

Consequently,

$$|J_{0,\delta}|<\frac{\delta}{\pi}\{|f_R'(x)|+1\}<\frac{\varepsilon}{3} \text{ for all } \delta<\frac{\varepsilon\pi}{3\{|f_R'(x)|+1\}} \qquad (11.216)$$

and for all values of l.

Next, let us estimate the integral

$$J_{\Delta,+\infty}=\frac{1}{\pi}\int_{\Delta}^{+\infty}f(x+\zeta)\frac{\sin l\zeta}{\zeta}\,d\zeta-\frac{f(x+0)}{\pi}\int_{\Delta}^{+\infty}\frac{\sin l\zeta}{\zeta}\,d\zeta$$

We can write

$$|J_{\Delta,+\infty}|\leqslant\frac{1}{\pi}\int_{\Delta}^{\infty}|f(x+\zeta)|\frac{d\zeta}{\zeta}+\frac{|f(x+0)|}{\pi}\left|\int_{\Delta}^{\infty}\frac{\sin l\zeta}{\zeta}\,d\zeta\right|\leqslant$$

$$\leqslant\frac{1}{\pi\Delta}\int_{-\infty}^{+\infty}|f(x+\zeta)|\,d\zeta+\frac{|f(x+0)|}{\pi}\left|\int_{l\Delta}^{+\infty}\frac{\sin\zeta^*}{\zeta^*}\,d\zeta^*\right|=$$

$$=\frac{Q}{\pi\Delta}+\frac{|f(x+0)|}{\pi}\left|\int_{l\Delta}^{+\infty}\frac{\sin\zeta^*}{\zeta^*}\,d\zeta^*\right| \text{ where } \zeta^*=l\zeta \qquad (11.217)$$

By condition (11.203), we have $Q=\int_{-\infty}^{+\infty}|f(x)|\,dx<\infty$ and, therefore, for all sufficiently large $\Delta>0$ the inequality $\frac{Q}{\pi\Delta}<\frac{\varepsilon}{6}$ is fulfilled for all l. Furthermore, the integral $\int_0^{+\infty}\frac{\sin\zeta^*}{\zeta^*}\,d\zeta^*$ being convergent, we have

$$\frac{|f(x+0)|}{\pi}\left|\int_{l\Delta}^{+\infty}\frac{\sin\zeta^*}{\zeta^*}\,d\zeta^*\right|<\frac{\varepsilon}{6}$$

for all sufficiently large $\Delta>0$ and all $l\geqslant 1$. Hence, by virtue of (11.217), the inequality

$$|J_{\Delta,+\infty}|<\frac{\varepsilon}{3} \qquad (11.218)$$

holds for all sufficiently large $\Delta>0$ and all $l\geqslant 1$.

Finally, let us estimate the integral

$$J_{\delta,\Delta}=\frac{1}{\pi}\int\limits_{\delta}^{\Delta}\frac{f(x+\zeta)-f(x+0)}{\zeta}\sin l\zeta\,d\zeta \qquad (11.219)$$

The expression $\dfrac{f(x+\zeta)-f(x+0)}{\zeta}$ is a piecewise smooth function of the argument ζ on the interval $\delta\leqslant\zeta\leqslant\Delta$. Consequently, by the key lemma (see § 5), the inequality

$$|J_{\delta,\Delta}|<\frac{\varepsilon}{3} \qquad (11.220)$$

is valid for all sufficiently large $l\geqslant 1$. On the basis of (11.216), (11.218) and (11.220) we conclude that, for all sufficiently large $l\geqslant 1$, we have the relation

$$|J_{0,+\infty}|<\varepsilon \qquad (11.221)$$

which is what we set out to prove.

Note. The above theorem on Fourier's integral can be proved under more general conditions imposed on the function $f(x)$. Namely, *if the function $f(x)$ is absolutely integrable over the x-axis and satisfies the conditions that (1) it is piecewise smooth on every finite interval of the x-axis and (2) the expression* $\left|\dfrac{f(x+\zeta)-f(x+0)}{\zeta}\right|$ *is bounded for any fixed x and all sufficiently small $\zeta>0$, the above theorem remains valid.*

Indeed, the proof of the theorem reduces to estimating the three integrals $J_{0,\delta}$, $J_{\delta,\Delta}$, and $J_{\Delta,+\infty}$ for $J_{0,+\infty}$ and the three integrals $J_{0,-\delta}$, $J_{-\delta,-\Delta}$ and $J_{-\Delta,-\infty}$ which are considered similarly. By the absolute integrability of $f(x)$, the integral $J_{0,\delta}$ is small for all sufficiently large Δ. The integral $J_{0,\delta}$ is small for all sufficiently small $\delta>0$ provided that the expression $\left|\dfrac{f(x+\zeta)-f(x+0)}{\zeta}\right|$ is bounded for every fixed x and all sufficiently small $\zeta>0$. To estimate the integral

$$J_{\delta,\Delta}=\frac{1}{\pi}\int\limits_{\delta}^{\Delta}\frac{f(x+\zeta)-f(x+0)}{\zeta}\sin l\zeta\,d\zeta$$

we note that the function $\varphi(\zeta)=\dfrac{f(x+\zeta)-f(x+0)}{\zeta}$ is piecewise continuous on the interval $0<\delta\leqslant\zeta\leqslant\Delta$ for any fixed x. Let $[a, b]$ be an interval on which $\varphi(\zeta)$ is continuous. Take an arbitrary $\varepsilon>0$. Let us construct a piecewise smooth function $g_\varepsilon(x)$ such that the inequality

$$|\varphi(\zeta)-g_\varepsilon(\zeta)|<\frac{\varepsilon}{2(b-a)} \quad \text{for } 0<\delta\leqslant\zeta\leqslant\Delta$$

holds (which can be done as in the proof of Weierstrass' trigonometric approximation theorem). Then we have

$$\left|\int_a^b \varphi(\zeta)\sin l\zeta\, d\zeta\right| \leqslant \int_a^b |\varphi(\zeta)-g_\varepsilon(\zeta)|\, d\zeta +$$

$$+\left|\int_a^b g_\varepsilon(\zeta)\sin l\zeta\, d\zeta\right| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

for all sufficiently large $l > 0$ since the key lemma is valid for the piecewise smooth function $g_\varepsilon(\zeta)$. Breaking up the integral $J_{\delta,\Delta}$ into the sum of integrals taken over the intervals of continuity of $\varphi(\zeta)$ we see that $J_{\delta,\Delta} \to 0$ for $l \to +\infty$, which completes the proof of the theorem.

3. Fourier Integral as an Expansion into a Sum of Harmonics. Fourier's integral formula (11.206) can be rewritten as

$$f(x) = \int_0^{+\infty} [A(\lambda)\cos\lambda x + B(\lambda)\sin\lambda x]\, d\lambda \tag{11.222}$$

where

$$A(\lambda) = \frac{1}{\pi}\int_{-\infty}^{+\infty} f(\xi)\cos\lambda\xi\, d\xi, \quad B(\lambda) = \frac{1}{\pi}\int_{-\infty}^{+\infty} f(\xi)\sin\lambda\xi\, d\xi \tag{11.223}$$

Relation (11.222) is analogous to an expansion of a function into a trigonometric series, and expressions (11.223) are similar to the formulas for the Fourier coefficients. Let us transform expression (11.222). We have

$$A(\lambda)\cos\lambda x + B(\lambda)\sin\lambda x = N(\lambda)\sin(\lambda x + \varphi_\lambda) \tag{11.224}$$

where

$$N(\lambda) = \sqrt{A^2(\lambda)+B^2(\lambda)}, \quad \cos\varphi_\lambda = \frac{A(\lambda)}{N(\lambda)}, \quad \sin\varphi_\lambda = \frac{B(\lambda)}{N(\lambda)} \tag{11.225}$$

Thus, relation (11.222) can be interpreted as an expansion of a function $f(x)$ defined for all x, $-\infty < x < +\infty$, into a sum of harmonic oscillations whose frequencies λ continuously cover the semi-infinite interval $0 \leqslant \lambda < +\infty$. The functions $A(\lambda)$ and $B(\lambda)$ (see (11.225)) give us the law of distribution of the amplitudes and initial phases φ_λ when λ varies over the positive half-axis λ, $0 \leqslant \lambda < +\infty$.

If a function $f(x)$ is defined on a finite interval $[-l, l]$ it can be expanded, under the conditions stated above, into harmonic

oscillations:

$$f(x)=\frac{a_0}{2}+\sum_{k=1}^{+\infty}\left(a_k\cos\frac{k\pi x}{l}+b_k\sin\frac{k\pi x}{l}\right)=$$

$$=\frac{a_0}{2}+\sum_{k=1}^{+\infty}N_k\sin(\lambda_k x+\varphi_k) \qquad (11.226)$$

where the frequencies $\lambda_k=\frac{k\pi}{l}$, $k=1, 2, \ldots$, form an arithmetical progression.

4. Fourier Integral in Complex Form. The Fourier integral formula can be rewritten in the complex form

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\,e^{i\lambda(x-\xi)}\,d\xi \qquad (11.227)$$

equivalent to (11.206): the integral $\int\limits_{-\infty}^{+\infty}f(\xi)\cos\lambda(x-\xi)\,d\xi$ is an even function of λ and the integral $\int\limits_{-\infty}^{+\infty}f(\xi)\sin\lambda(x-\xi)\,d\xi$ is an odd function of λ, and therefore, we have

$$\frac{1}{\pi}\int_{0}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\cos\lambda(\xi-x)\,d\xi=\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\cos\lambda(x-\xi)\,d\xi$$

and

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\sin\lambda(x-\xi)\,d\xi=0$$

Consequently, by Euler's formula

$$e^{i\lambda(x-\xi)}\equiv\cos\lambda(x-\xi)+i\sin\lambda(x-\xi)$$

we can write

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\,e^{i\lambda(x-\xi)}\,d\xi=\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\cos\lambda(x-\xi)\,d\xi+$$

$$+\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\sin\lambda(x-\xi)\,d\xi=$$

$$=\frac{1}{\pi}\int_{0}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\cos\lambda(x-\xi)\,d\xi$$

whence it follows that formulas (11.206) and (11.227) are equivalent. But it should be noted that in the general case the integral $\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} d\lambda \int\limits_{-\infty}^{+\infty} f(\xi)\sin\lambda(x-\xi)\,d\xi = 0$ entering into the relation is understood in the sense of Cauchy's principal value (see § 3 of Chapter 9), that is

$$\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} d\lambda \int\limits_{-\infty}^{+\infty} f(\xi)\sin\lambda(x-\xi)\,d\xi =$$

$$= \lim_{l\to+\infty}\frac{1}{2\pi}\int\limits_{-l}^{l} d\lambda \int\limits_{-\infty}^{+\infty} f(\xi)\sin\lambda(x-\xi)\,d\xi = 0$$

5. Fourier Transformation. Equality (11.227) can be rewritten in the form

$$f(x) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{i\lambda x}\,d\lambda\left(\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} f(\xi)\,e^{-i\lambda\xi}\,d\xi\right) \qquad (11.228)$$

Introducing the notation

$$\overline{f}(\lambda) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} f(\xi)\,e^{-i\lambda\xi}\,d\xi \qquad (11.229)$$

we obtain, by (11.228), the formula

$$f(x) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} \overline{f}(\lambda)\,e^{i\lambda x}\,d\lambda \qquad (11.230)$$

The function $\overline{f}(\lambda)$ is called the **Fourier transform (or spectral characteristic) of the function** $f(x)$ defined on the real x-axis, $-\infty < x < +\infty$. The transformation from $f(x)$ to $\overline{f}(\lambda)$ performed according to formula (11.229) is called the **Fourier transformation.** Formula (11.230) (**Fourier's inversion formula**) expressing the original function $f(x)$ in terms of its Fourier transform $\overline{f}(\lambda)$ describes the **(Fourier) inverse transformation,** and $f(x)$ is termed the **Fourier inverse transform** of $\overline{f}(\lambda)$.

Now we can rephrase Theorem 11.11 on Fourier's integral as follows:

Theorem 11.12. *If $f(x)$ is an absolutely integrable function (on the whole x-axis) piecewise smooth on every finite interval of the x-axis, then* (1) *the Fourier transform determined by formula* (11.229) *exists, and* (2) *we have inversion formula* (11.230) *which should be*

understood as the limiting relation

$$f(x)=\frac{1}{\sqrt{2\pi}}\lim_{l\to+\infty}\int_{-l}^{l}\bar{f}(\lambda)\,e^{i\lambda x}\,d\lambda$$

Note. According to the note after Theorem 11.11, we can assert that Theorem 11.12 remains valid for every function $f(x)$ which is absolutely integrable over the entire x-axis and piecewise continuous on every finite interval of the x-axis provided the expression $\left|\frac{f(x+\zeta)-f(x+0)}{\zeta}\right|$ is bounded for every fixed x and all sufficiently small $|\zeta|\neq 0$. Fourier transforms of functions defined for $-\infty<x<+\infty$ are widely applied to various problems of mathematics and mathematical physics (see Appendix 4 to Chapter 11).

For functions defined over the semi-infinite interval $0\leqslant x<+\infty$ we also use the so-called **Fourier sine transform** and **Fourier cosine transform.** Let us dwell in more detail on these notions. Applying the formula

$$\cos\lambda(\xi-x)=\cos\lambda\xi\cos\lambda x+\sin\lambda\xi\sin\lambda x$$

to the integrand in formula (11.206) we deduce

$$f(x)=\frac{1}{\pi}\int_{0}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\cos\lambda\xi\cos\lambda x\,d\xi+$$
$$+\frac{1}{\pi}\int_{0}^{+\infty}d\lambda\int_{-\infty}^{+\infty}f(\xi)\sin\lambda\xi\sin\lambda x\,d\xi \qquad (11.231)$$

where, by the absolute integrability of $f(x)$ over the whole x-axis, both integrals are convergent. If $f(\xi)$ is an even function the product $f(\xi)\sin\lambda\xi$ is an odd function while the product $f(\xi)\cos\lambda\xi$ is an even function. Therefore the second term on the right-hand side of (11.231) turns into zero and we thus obtain

$$f(x)=\frac{2}{\pi}\int_{0}^{+\infty}\cos\lambda x\,d\lambda\int_{0}^{+\infty}f(\xi)\cos\lambda\xi\,d\xi \qquad (11.232)$$

Similarly, if $f(x)$ is an odd function we find that

$$f(x)=\frac{2}{\pi}\int_{0}^{+\infty}\sin\lambda x\,d\lambda\int_{0}^{+\infty}f(\xi)\sin\lambda\xi\,d\xi \qquad (11.233)$$

If x is a point of discontinuity of the function $f(x)$ the left-hand sides of equalities (11.231), (11.232) and (11.233) should be replaced by the expression $\frac{f(x+0)+f(x-0)}{2}$.

Now suppose that a function $f(x)$ is defined only on the interval $0 \leqslant x < +\infty$. Then it can be extended to the negative half of x-axis in such a way that we obtain an even or an odd function defined for all x, $-\infty < x < +\infty$. If $f(x)$ is extended in even fashion we obtain the representation

$$f(x)=\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}\cos\lambda x\,d\lambda\left(\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}f(\xi)\cos\lambda\xi\,d\xi\right) \qquad (11.234)$$

and if it is extended in odd fashion we get another representation:

$$f(x)=\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}\sin\lambda x\,d\lambda\left(\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}f(\xi)\sin\lambda\xi\,d\xi\right) \qquad (11.235)$$

If the function $f(x)$ defined on the semi-axis $0 \leqslant x < +\infty$ is continuous at the point $x = 0$ then, after it has been extended as an even function to the entire x-axis, it remains continuous at the point $x = 0$, and therefore relation (11.234) will also hold for $x = 0$ in this case. On the contrary, for the function $f(x)$ extended in odd fashion relation (11.235) does not hold for $x = 0$ in the general case even if the original function is continuous at the point $x = 0$. It is clear that this relation may only hold for $x = 0$ if $f(0) = 0$ since we have the equality $\frac{f(+0)+f(-0)}{2}=0$ for every odd function.

Equality (11.234) can be rewritten in another form. Putting

$$f_c(\lambda)=\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}f(\xi)\cos\lambda\xi\,d\xi \qquad (11.236)$$

we derive from (11.234) the formula

$$f(x)=\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}f_c(\lambda)\cos\lambda x\,d\lambda \qquad (11.237)$$

which is equivalent to (11.234). The function $f_c(\lambda)$ is called the **Fourier cosine transform** of the function $f(x)$ defined on the semi-axis $0 \leqslant x < +\infty$. Accordingly, the transformation from $f(x)$ to $f_c(\lambda)$ performed by formula (11.236) is termed the **Fourier cosine transformation**. The **Fourier inverse cosine transformation** described by formula (11.237) yields the expression of $f(x)$ (which is **Fourier's inverse cosine transform** of $f_c(\lambda)$) in terms of the function $f_c(\lambda)$.

We see that transformations (11.236) and (11.237) are inverse to each other.

Similarly, instead of relation (11.235) we can write the formulas

$$f_s(\lambda)=\sqrt{\frac{2}{\pi}}\int_0^{+\infty} f(\xi)\sin\lambda\xi\,d\xi \tag{11.238}$$

and

$$f(x)=\sqrt{\frac{2}{\pi}}\int_0^{+\infty} f_s(\lambda)\sin\lambda x\,d\lambda \tag{11.239}$$

which are equivalent to (11.235). The function $f_s(\lambda)$ is known as the Fourier sine transform of the function $f(x)$ defined on the semi-infinite interval $0\leqslant x<+\infty$, and transformation (11.238) from $f(x)$ to $f_s(\lambda)$ is called the **Fourier sine transformation**. Formula (11.239) describes **Fourier's inverse sine transformation** which expresses the original function $f(x)$ (**Fourier's inverse sine transform** of $f_s(\lambda)$) in terms of $f_s(\lambda)$.

Examples

1. Consider the function

$$f(x)=\begin{cases}1 & \text{for } 0\leqslant x<a\\ \frac{1}{2} & \text{for } x=a\\ 0 & \text{for } x>a\end{cases}$$

Its Fourier cosine transform is the function

$$f_c(\lambda)=\sqrt{\frac{2}{\pi}}\int_0^{a}\cos\lambda\xi\,d\xi=\sqrt{\frac{2}{\pi}}\,\frac{\sin\lambda x}{\lambda}$$

Applying formula (11.237) we obtain

$$\frac{2}{\pi}\int_0^{+\infty}\frac{\sin\lambda a\cos\lambda x}{\lambda}\,d\lambda=f(x)\qquad\begin{cases}1 & \text{for } 0\leqslant x<a\\ \frac{1}{2} & \text{for } x=a\\ 0 & \text{for } x>a\end{cases}$$

2. For the function $f(x)=e^{-ax}$, $a>0$, $x\geqslant 0$, we find, performing integration by parts in formulas (11.236) and (11.238) the expressions

$$f_c(\lambda)=\sqrt{\frac{2}{\pi}}\int_0^{+\infty}e^{-a\xi}\cos\lambda\xi\,d\xi=\sqrt{\frac{2}{\pi}}\,\frac{a}{a^2+\lambda^2}$$

and

$$f_s(\lambda)=\sqrt{\frac{2}{\pi}}\int_0^{+\infty}e^{-a\xi}\sin\lambda\xi\,d\xi=\sqrt{\frac{2}{\pi}}\,\frac{\lambda}{a^2+\lambda^2}$$

Accordingly, applying formulas (11.237) and (11.239) to the above equalities we obtain

$$\frac{2a}{\pi}\int_{0}^{+\infty}\frac{\cos\lambda x}{a^2+\lambda^2}\,d\lambda=e^{-ax},\quad x>0$$

and

$$\frac{2}{\pi}\int_{0}^{+\infty}\frac{\lambda\sin\lambda x}{a^2+\lambda^2}\,d\lambda=e^{-ax},\quad x>0$$

We see that Fourier's cosine and sine transformations enable us to find the values of some integrals dependent on a parameter. But the main application of these transformations lies in using them for solving various problems of mathematical physics (see Appendix 3 to Chapter 11).

6. Fourier Integral for Functions of Several Independent Variables. We shall begin with the case of two independent variables. Let a function $f(x_1, x_2)$ be defined for $-\infty < x_1 < +\infty$, $-\infty < x_2 < +\infty$. We shall suppose that $f(x_1, x_2)$ is absolutely integrable with respect to each variable x_1 and x_2 from $-\infty$ to $+\infty$ for every fixed value of the other variable. If, in addition, the function $f(x_1, x_2)$ is continuous and piecewise smooth with respect to x_1 (x_2) for every fixed value of x_2 (x_1) we can apply Fourier's integral formula to each variable for every fixed value of the other. Fixing an arbitrary value of x_2 and applying Fourier's formula (11.206) (for the variable x_1) we obtain

$$f(x_1, x_2)=\frac{1}{\pi}\int_{0}^{+\infty}d\lambda_1\int_{-\infty}^{+\infty}f(\xi_1, x_2)\cos\lambda_1(x_1-\xi_1)\,d\xi_1 \tag{11.240}$$

Similarly, we can fix $x_1=\xi_1$ and apply Fourier's formula (11.206) to the variable x_2 which results in

$$f(\xi_1, x_2)=\frac{1}{\pi}\int_{0}^{+\infty}d\lambda_2\int_{-\infty}^{+\infty}f(\xi_1, \xi_2)\cos\lambda_2(x_2-\xi_2)\,d\xi_2 \tag{11.241}$$

Substituting (11.241) into (11.240) we derive the formula

$$f(x_1, x_2)=\frac{1}{\pi^2}\int_{0}^{+\infty}d\lambda_1\int_{-\infty}^{+\infty}\cos\lambda_1(x_1-\xi_1)\,d\xi_1\int_{0}^{+\infty}d\lambda_2\int_{-\infty}^{+\infty}f(\xi_1, \xi_2)\times$$

$$\times\cos\lambda_2(x_2-\xi_2)\,d\xi_2=\frac{1}{\pi^2}\int_{0}^{+\infty}d\lambda_1\int_{-\infty}^{+\infty}d\xi_1\int_{0}^{+\infty}d\lambda_2\int_{-\infty}^{+\infty}f(\xi_1, \xi_2)\times$$

$$\times\cos\lambda_1(x_1-\xi_1)\cos\lambda_2(x_2-\xi_2)\,d\xi_2 \tag{11.242}$$

If $f(x_1, x_2)$ is an even function of each argument x_1 and x_2 formula (11.242) turns into

$$f(x_1, x_2) = \frac{4}{\pi^2} \int\limits_0^{+\infty} \cos \lambda_1 x_1 \, d\lambda_1 \int\limits_0^{+\infty} \cos \lambda_1 \xi_1 \, d\xi_1 \times$$

$$\times \int\limits_0^{+\infty} \cos \lambda_2 x_2 \, d\lambda_2 \int\limits_0^{+\infty} f(\xi_1, \xi_2) \cos \lambda_2 \xi_2 \, d\xi_2 \qquad (11.243)$$

Similarly, if $f(x_1, x_2)$ is an odd function of each argument x_1 and x_2 we obtain

$$f(x_1, x_2) = \frac{4}{\pi^2} \int\limits_0^{+\infty} \sin \lambda_1 x_1 \, d\lambda_1 \int\limits_0^{+\infty} \sin \lambda_1 \xi_1 \, d\xi_1 \times$$

$$\times \int\limits_0^{+\infty} \sin \lambda_2 x_2 \, d\lambda_2 \int\limits_0^{+\infty} f(\xi_1, \xi_2) \sin \lambda_2 \xi_2 \, d\xi_2 \qquad (11.244)$$

Passing to the complex form of Fourier's integral we can rewrite formula (11.242) in the form

$$f(x_1, x_2) =$$

$$= \frac{1}{(2\pi)^3} \int\limits_{-\infty}^{+\infty} d\lambda_1 \int\limits_{-\infty}^{+\infty} d\xi_1 \int\limits_{-\infty}^{+\infty} d\lambda_2 \int\limits_{-\infty}^{+\infty} f(\xi_1, \xi_2) \, e^{i[\lambda_1(x_1-\xi_1)+\lambda_2(x_2-\xi_2)]} \, d\xi_2 \quad (11.245)$$

where the integrals taken with respect to λ_1 and λ_2 should be understood, in the general case, in the sense of Cauchy's principal value (see § 3 of Chapter 9 and § 9, Sec. 4 of the present chapter). If it is allowable to reverse the order of integration with respect to ξ_1 and λ_2, formula (11.245) turns to be equivalent to the following two formulas:

$$\bar{\bar{f}}(\lambda_1, \lambda_2) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} d\xi_1 \int\limits_{-\infty}^{+\infty} f(\xi_1, \xi_2) \, e^{-i[\lambda_1\xi_1+\lambda_2\xi_2]} \, d\xi_2 \qquad (11.246)$$

and

$$f(x_1, x_2) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} d\lambda_1 \int\limits_{-\infty}^{+\infty} \bar{\bar{f}}(\lambda_1, \lambda_2) \, e^{i[\lambda_1 x_1+\lambda_2 x_2]} \, d\lambda_2 \qquad (11.247)$$

Formula (11.246) expresses the Fourier transformation from $f(x_1, x_2)$ to $f(\lambda_1, \lambda_2)$ and formula (11.247) describes Fourier's inverse transformation. Accordingly, $f(\lambda_1, \lambda_2)$ is the (two-dimensional) Fourier transform of $f(x_1, x_2)$, and $f(x_1, x_2)$ is the Fourier inverse transform of $\bar{\bar{f}}(\lambda_1, \lambda_2)$.

In the case of three or more independent variables the Fourier transformations are constructed in a similar manner. Here we shall give the corresponding formulas for functions of three independent variables. The Fourier integral formula is written as

$$f(x_1, x_2, x_3) = \frac{1}{\pi^3} \int_0^{+\infty} d\lambda_1 \int_{-\infty}^{+\infty} d\xi_1 \int_0^{+\infty} d\lambda_2 \int_{-\infty}^{+\infty} d\xi_2 \int_0^{+\infty} d\lambda_3 \int_{-\infty}^{+\infty} f(\xi_1, \xi_2, \xi_3) \times$$
$$\times \cos\lambda_1(x_1-\xi_1)\cos\lambda_2(x_2-\xi_2)\cos\lambda_3(x_3-\xi_3)\,d\xi_3 \qquad (11.248)$$

or

$$f(x_1, x_2, x_3) = \frac{1}{(2\pi)^3} \int_0^{+\infty} d\lambda_1 \int_{-\infty}^{+\infty} d\xi_1 \int_0^{+\infty} d\lambda_2 \int_{-\infty}^{+\infty} d\xi_2 \int_{-\infty}^{+\infty} d\lambda_3 \int_{-\infty}^{+\infty} f(\xi_1, \xi_2, \xi_3) \times$$
$$\times e^{i[\lambda_1(x_1-\xi_1)+\lambda_2(x_2-\xi_2)+\lambda_3(x_3-\xi_3)]}\,d\xi_3 \qquad (11.249)$$

in complex form.

If the conditions which guarantee the possibility of reversing the order of integration are fulfilled, formula (11.249) is equivalent to the following two formulas:

$$\overline{\overline{\overline{f}}}(\lambda_1, \lambda_2, \lambda_3) = \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{+\infty} d\xi_1 \int_{-\infty}^{+\infty} d\xi_2 \int_{-\infty}^{+\infty} f(\xi_1, \xi_2, \xi_3) \times$$
$$\times e^{-i[\lambda_1\xi_1+\lambda_2\xi_2+\lambda_3\xi_3]}\,d\xi_3 \qquad (11.250)$$

and

$$f(x_1, x_2, x_3) = \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{+\infty} d\lambda_1 \int_{-\infty}^{+\infty} d\lambda_2 \int_{-\infty}^{+\infty} \overline{\overline{\overline{f}}}(\lambda_1, \lambda_2, \lambda_3) \times$$
$$\times e^{i[\lambda_1x_1+\lambda_2x_2+\lambda_3x_3]}\,d\lambda_3 \qquad (11.251)$$

In the general case the integrals on the right-hand sides of (11.247) and (11.251) are understood in the sense of Cauchy's principal value.

We now briefly discuss the justification of formulas (11.246) and (11.247). Formulas (11.250) and (11.251) are proved in a similar way.

Theorem 11.13. Let a function $f(x_1, x_2)$ be continuous throughout the x_1, x_2-plane. Suppose that the following conditions are fulfilled:

(1) *The integrals*

$$\int_{-\infty}^{+\infty} |f(x_1, x_2)|\,dx_1 \quad and \quad \int_{-\infty}^{+\infty} |f(x_1, x_2)|\,dx_2 \qquad (11.252)$$

are uniformly convergent with respect to x_2 and x_1 on every finite interval $\underline{x}_2 \leqslant x_2 \leqslant \overline{x}_2$ and $\underline{x}_1 \leqslant x_1 \leqslant \overline{x}_1$.

(2) *The iterated integral*

$$\int_{-\infty}^{+\infty} dx_2 \int_{-\infty}^{+\infty} |f(x_1, x_2)|\, dx_1 \tag{11.253}$$

is convergent.

(3) *For all sufficiently small* $|\zeta| \neq 0$ *the inequality*

$$\left| \frac{f(x_1-\zeta, x_2) - f(x_1+0, x_2)}{\zeta} \right| \leqslant C_1 = \text{const}$$

holds for every fixed x_1 *and all* x_2.

(4) *For all sufficiently small* $|\zeta| \neq 0$ *the relation*

$$\left| \frac{f(x_1, x_2+\zeta) - f(x_1, x_2+0)}{\zeta} \right| \leqslant C_2(x_1)$$

is fulfilled for every fixed x_2 *and all* x_1 *where* $C_2(x_1)$ *is a function such that the integral* $\int\limits_{-\infty}^{+\infty} C_2(x_1)\, dx_1$ *converges.*

Then the two-dimensional Fourier transform

$$\overline{\overline{f}}(\lambda_1, \lambda_2) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\xi_1 \int_{-\infty}^{+\infty} f(\xi_1, \xi_2)\, e^{-i[\lambda_1\xi_1+\lambda_2\xi_2]}\, d\xi_2 \tag{11.254}$$

of the function $f(x_1, x_2)$ *exists, and we have the* ***Fourier inversion formula***

$$f(x_1, x_2) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\lambda_1 \int_{-\infty}^{+\infty} \overline{\overline{f}}(\lambda_1, \lambda_2)\, e^{i[\lambda_1x_1+\lambda_2x_2]}\, d\lambda_2 \tag{11.255}$$

which is understood in the sense of the limiting relation

$$f(x_1, x_2) = \frac{1}{2\pi} \lim_{l_1, l_2 \to +\infty} \int_{-l_1}^{l_1} d\lambda_1 \int_{-l_2}^{l_2} \overline{\overline{f}}(\lambda_1, \lambda_2)\, e^{i[\lambda_1x_1+\lambda_2x_2]}\, d\lambda_2 \tag{11.256}$$

where we first pass to the limit for $l_2 \to +\infty$ *and then for* $l_1 \to +\infty$.

Proof. The one-dimensional Fourier transform (with respect to the argument x_1)

$$\overline{f}(\lambda_1, x_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(\xi_1, x_2)\, e^{-i\lambda_1\xi_1}\, d\xi_1 \tag{11.257}$$

exists since the first integral (11.252) is convergent. The uniform convergence of this integral implies that the integral on the right-hand side of (11.257) is uniformly convergent and therefore the function $\overline{f}(\lambda_1, x_2)$ is continuous in x_2. By the hypothesis, inte-

gral (11.253) converges and therefore the integral

$$\int_{-\infty}^{+\infty} |\overline{f}(\lambda_1, x_2)|\, dx_2 = \int_{-\infty}^{+\infty} dx_2 \left| \int_{-\infty}^{+\infty} f(\xi_1, x_2)\, e^{i\lambda_1\xi_1}\, d\xi_1 \right| \frac{1}{\sqrt{2\pi}} \qquad (11.258)$$

also converges.

Taking into account conditions (1) and (3), the continuity of $f(x_1, x_2)$ as a function of x_1 and the note after Theorem 11.10 (see Sec. 4 of § 6), we see that the inversion formula

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \overline{f}(\lambda_1, x_2)\, e^{i\lambda_1 x_1}\, d\lambda_1 =$$

$$= \frac{1}{\sqrt{2\pi}} \lim_{l_1 \to +\infty} \int_{-l_1}^{l_1} \overline{f}(\lambda_1, x_2)\, e^{i\lambda_1 x_1}\, d\lambda_1 \qquad (11.259)$$

is valid here. By condition (4) and equality (11.257), we can write

$$|\overline{f}(\lambda_1, x_2+\zeta) - \overline{f}(\lambda_1, x_2+0)| \leqslant$$

$$\leqslant \int_{-\infty}^{+\infty} |f(\xi_1, x_2+\zeta) - f(\xi_1, x_2+0)|\, d\xi_1 \leqslant |\zeta| \int_{-\infty}^{+\infty} C_2(\xi_1)\, d\xi_1$$

that is

$$\left| \frac{\overline{f}(\lambda_1, x_2+\zeta) - \overline{f}(\lambda_1, x_2+0)}{\zeta} \right| < \int_{\infty}^{+\infty} C_2(\xi_1)\, d\xi_1 \qquad (11.260)$$

Since integral (11.258) is convergent the two-dimensional Fourier transform

$$\overline{\overline{f}}(\lambda_1, \lambda_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \overline{f}(\lambda_1, \xi_2)\, e^{-i\lambda_2\xi_2}\, d\xi_2 =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\xi_2 \int_{-\infty}^{+\infty} f(\xi_1, \xi_2)\, e^{-i[\lambda_1\xi_1 + \lambda_2\xi_2]}\, d\xi_1 \qquad (11.261)$$

of the function $f(x_1, x_2)$ exists, and the inversion formula

$$\overline{f}(\lambda_1, x_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \overline{\overline{f}}(\lambda_1, \lambda_2)\, e^{i\lambda_2 x_2}\, d\lambda_2 =$$

$$= \frac{1}{\sqrt{2\pi}} \lim_{l_2 \to +\infty} \int_{-l_2}^{l_2} \overline{\overline{f}}(\lambda_1, \lambda_2)\, e^{i\lambda_2 x_2}\, d\lambda_2 \qquad (11.262)$$

is valid because the function $\overline{f}(\lambda_1, x_2)$ is continuous in x_2 and condition (11.260) is fulfilled. Substituting (11.262) into (11.259) we

obtain

$$f(x_1, x_2) = \frac{1}{2\pi} \lim_{l_1 \to +\infty} \int_{-l_1}^{l_1} e^{i\lambda_1 x_1} d\lambda_1 \left\{ \lim_{l_2 \to +\infty} \int_{-l_2}^{l_2} \overline{\overline{f}} (\lambda_1, \lambda_2) e^{i\lambda_2 x_2} d\lambda_2 \right\} \tag{11.263}$$

or, which is the same,

$$f(x_1, x_2) = \frac{1}{2\pi} \lim_{l_1, l_2 \to +\infty} \int_{-l_1}^{l_1} d\lambda_1 \int_{-l_2}^{l_2} \overline{\overline{f}} (\lambda_1, \lambda_2) e^{i[\lambda_1 x_1 + \lambda_2 x_2]} d\lambda_2 \tag{11.264}$$

where the passage to the limit is first performed with respect to l_2 and then to l_1. The theorem has thus been proved.

In the case of three independent variables the formulas for the **three-dimensional Fourier transform** and the corresponding **inverse transform** are written in the form

$$\overline{\overline{\overline{f}}} (\lambda_1, \lambda_2, \lambda_3) = \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi_1, \xi_2, \xi_3) \times$$
$$\times e^{-i[\lambda_1 \xi_1 + \lambda_2 \xi_2 + \lambda_3 \xi_3]} d\xi_1 \, d\xi_2 \, d\xi_3 \tag{11.265}$$

and

$$f(x_1, x_2, x_3) = \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \overline{\overline{\overline{f}}} (\lambda_1, \lambda_2, \lambda_3) \times$$
$$\times e^{i[\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3]} d\lambda_1 \, d\lambda_2 \, d\lambda_3 \tag{11.266}$$

Substituting (11.265) into (11.266) we arrive at the relation

$$f(x_1, x_2, x_3) = \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi_1, \xi_2, \xi_3) \times$$
$$\times e^{i[\lambda_1 (x_1 - \xi_1) + \lambda_2 (x_2 - \xi_2) + \lambda_3 (x_3 - \xi_3)]} d\xi_1 \, d\xi_2 \, d\xi_3 \, d\lambda_1 \, d\lambda_2 \, d\lambda_3 \tag{11.267}$$

In the general case of N ($N > 3$) independent variables the corresponding formulas for Fourier's transforms can be easily written down by analogy with the above formulas.

To justify equalities (11.265) and (11.266) we should introduce the condition that the corresponding integrals (analogous to those entering into the formulation of the foregoing theorem) are convergent and, in addition, impose the requirement that for all sufficiently small $|\zeta| \neq 0$ the following inequalities hold:

(1) $$\left| \frac{f(x_1 + \zeta, x_2, x_3) - f(x_1 + 0, x_2, x_3)}{\zeta} \right| \leqslant C_1 = \text{const}$$

for every fixed x_1 and all x_2 and x_3,

(2) $$\left| \frac{f(x_1, x_2 + \zeta, x_3) - f(x_1, x_2 + 0, x_3)}{\zeta} \right| \leqslant C_2(x_1)$$

for every fixed x_2 and all x_1 and x_3,

$$(3)\quad \left|\frac{f(x_1, x_2, x_3+\zeta)-f(x_1, x_2, x_3+0)}{\zeta}\right| \leqslant C_3(x_1, x_2)$$

for every fixed x_3 and all x_1 and x_2 where $C_2(x_1)$ and $C_3(x_1, x_2)$ are some functions such that the integrals $\int_{-\infty}^{+\infty} C_2(x_1)\,dx_1$ and $\int_{-\infty}^{+\infty} dx_2 \int_{-\infty}^{+\infty} C_3(x_1, x_2)\,dx_1$ are convergent. Then, under these restrictions, the three-dimensional Fourier transform of the function $f(x_1, x_2, x_3)$ determined by formula (11.265) exists and equality (11.266) holds, the latter being understood in the sense of the limiting relation

$$f(x_1, x_2, x_3)=\frac{1}{(\sqrt{2\pi})^3}\lim_{l_1\to+\infty}\int_{-l_1}^{l_1} d\lambda_1 \left\{\lim_{l_2\to+\infty}\int_{-l_2}^{l_2} d\lambda_2 \times\right.$$

$$\left.\times\left[\lim_{l_3\to+\infty}\int_{-l_3}^{l_3} \overline{\overline{f}}(\lambda_1, \lambda_2, \lambda_3)\, e^{i[\lambda_1x_1+\lambda_2x_2+\lambda_3x_3]}\,d\lambda_3\right]\right\} \quad (11.268)$$

where the passage to the limit is first performed with respect to l_3, then to l_2 and finally with respect to l_1.

The general case of N independent variables ($N > 3$) is treated similarly.

APPENDIX 1 TO CHAPTER 11

ON LEGENDRE'S POLYNOMIALS

Here we shall prove that Legendre's polynomials

$$P_0(x)=1,\quad P_n(x)=\frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^n],\qquad n=1, 2, \ldots \quad (1)$$

are orthogonal on the interval $[-1, 1]$, i.e.

$$\int_{-1}^{1} P_n(x)\,P_m(x)\,dx=0 \qquad \text{for } m\neq n \quad (2)$$

Since m and n are involved equivalently in (2) it is sufficient to prove that relation (2) holds for $m < n$. For this purpose we must only show that

$$\int_{-1}^{1} P_n(x)\,x^m\,dx=0 \quad \text{for } m<n \quad (3)$$

where m is a nonnegative integer. Putting

$$P_n(x)=\frac{1}{2^n n!}\frac{d^n u_n(x)}{dx^n}, \qquad u_n(x)=[x^2-1]^n$$

we can write

$$\int_{-1}^{1} P_n(x)\,x^m\,dx=\frac{1}{2^n n!}\int_{-1}^{1}\frac{d^n u_n(x)}{dx^n}\,x^m\,dx \tag{4}$$

Integrating by parts in (4) $m+1$ times and taking into account that

$$u_n(\pm 1)=u_n'(\pm 1)=\ldots=u_n^{(n-1)}(\pm 1)=0$$

we arrive at equality (3) and thus the orthogonality of Legendre's polynomials on the interval $[-1, 1]$ has been established.

Now let us compute the norm of the nth polynomial $P_n(x)$. For this purpose we integrate by parts in the integral

$$\|P_n(x)\|^2=\frac{1}{2^n n!}\int_{-1}^{1}\left[\frac{d^n u_n(x)}{dx^n}\right]^2 dx \tag{5}$$

Integrating by parts n times and taking into account that $u_n(x)$ is a polynomial of degree $2n$ and $u_n(\pm 1)=u_n'(\pm 1)=\ldots$ $\ldots=u_n^{(n-1)}(\pm 1)=0$ we obtain

$$\int_{-1}^{1}\frac{d^n u_n(x)}{dx^n}\frac{d^n u_n(x)}{dx^n}\,dx=-\int_{-1}^{1}\frac{d^{n-1}u_n(x)}{dx^{n-1}}\frac{d^{n+1}u_n(x)}{dx^{n+1}}\,dx=$$

$$=(-1)^2\int_{-1}^{1}\frac{d^{(n-2)}u_n(x)}{dx^{n-2}}\frac{d^{n+2}u_n(x)}{dx^{n+2}}\,dx=\ldots=$$

$$=(-1)^n\int_{-1}^{1}u_n(x)\frac{d^{2n}u_n(x)}{dx^{2n}}\,dx=$$

$$=(2n)!\int_{-1}^{1}(1-x)^n(1+x)^n\,dx \tag{6}$$

But we have

$$\int_{-1}^{1}(1-x)^n(1+x)^n\,dx=\frac{n}{n+1}\int_{-1}^{1}(1-x)^{n-1}(1+x)^{n+1}\,dx=\ldots=$$

$$=\frac{n(n-1)\ldots 1}{(n+1)(n+2)\ldots(2n)}\int_{-1}^{1}(1+x)^{2n}\,dx=$$

$$=\frac{(n!)^2}{(2n)!\,(2n+1)}2^{2n+1} \tag{7}$$

and therefore, substituting (6) and (7) into (5), we get the formula

$$\| P_n(x) \|^2 = \int_{-1}^{1} [P_n(x)]^2\,dx = \frac{2}{2n+1} \tag{8}$$

Consequently, the norm of the nth polynomial $P_n(x)$ is expressed as

$$\| P_n(x) \| = \sqrt{\frac{2}{2n+1}} \tag{9}$$

It should be noted that the nth polynomial $P_n(x)$ is of degree n for $n = 0, 1, \ldots$. Legendre's polynomials $P_0(x), P_1(x), \ldots, P_n(x)$ are orthogonal on the interval $[-1, 1]$ and hence they are linearly independent. Consequently, the system of Legendre's polynomials is a basis of the space of all algebraic polynomials of degree not greater than n. It follows that every polynomial of degree not greater than n can be represented in the form of a linear combination of Legendre's polynomials $P_0(x), P_1(x), \ldots, P_n(x)$. In particular, we have

$$x^n = \alpha_{0n} P_0(x) + \alpha_{1n} P_1(x) + \ldots + \alpha_{nn} P_n(x)$$

(see Appendix 2 to Chapter 11).

APPENDIX 2 TO CHAPTER 11

ORTHOGONALITY WITH WEIGHT FUNCTION AND ORTHOGONALIZATION PROCESS

The concept of *orthogonality with weight function* is a generalization of the concept of orthogonality of functions in the sense of relation (11.68).

Let $p(x)$ be a nonnegative function which is not identically equal to zero. We shall suppose that this function is continuous in an open interval (a, b) and that the integral

$$\int_a^b p(x)\,dx \tag{1}$$

exists (as proper or improper integral) and is positive.* The function $p(x)$ will be referred to as a weight function.

* This integral may turn out to be improper under these conditions if the function $p(x)$ is unbounded for $x \to a+0$ or $x \to b-0$. We encounter singularities of $p(x)$ of this type when studying some important classes of *special functions* (e.g. such a weight function is used at the end of this appendix when we consider Chebyshev's polynomials).

Let $f(x)$ be a function defined on $[a, b]$ such that the integrals

$$\int_a^b p(x) f(x)\,dx \quad \text{and} \quad \int_a^b p(x)[f(x)]^2\,dx \tag{2}$$

exist (as proper or improper integrals). Then the function $f(x)$ is said to be **square-integrable with weight function** $p(x)$ **on the interval** $[a, b]$. In particular, if $p(x) \equiv 1$ we come back to the ordinary definition of a square-integrable function given in Sec. 3 of Chapter 11 (see relation (11.93)).

Let

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots \tag{3}$$

be a system of functions defined on $[a, b]$ which are square-integrable with weight function $p(x)$ on $[a, b]$, that is the integrals

$$\int_a^b p(x)\varphi_n(x)\,dx \quad \text{and} \quad \int_a^b p(x)[\varphi_n(x)]^2\,dx, \quad n = 1, 2, \ldots \tag{4}$$

(understood as proper or improper integrals) exist. If the interval $[a, b]$ is finite the existence of integrals (2) and (4) and the obvious inequalities

$$|f(x)\varphi_n(x)| \leqslant \frac{1}{2}[f^2(x) + \varphi_n^2(x)]$$

and (5)

$$|\varphi_n(x)\varphi_m(x)| \leqslant \frac{1}{2}[\varphi_n^2(x) + \varphi_m^2(x)]$$

imply that the integrals

$$\int_a^b p(x) f(x)\varphi_n(x)\,dx \quad \text{and} \quad \int_a^b p(x)\varphi_n(x)\varphi_m(x)\,dx \tag{6}$$

also exist. Let us agree that if the interval $[a, b]$ is infinite we additionally impose the requirement that integrals (6) exist.

In what follows we shall suppose that every function we deal with is continuous everywhere on $[a, b]$ except possibly at a finite number of points which, in particular, may be singular points of the functions.

We say that two functions $\varphi_n(x)$ and $\varphi_m(x)$ are ***orthogonal on the interval*** $[a, b]$ ***with weight function*** $p(x)$ *if*

$$\int_a^b p(x)\varphi_n(x)\varphi_m(x)\,dx = 0 \qquad \textit{for } m \neq n \tag{7}$$

System (2) of functions square-integrable with weight function $p(x)$ on the interval $[a, b]$ is said to be ***orthogonal*** *on $[a, b]$* ***with***

weight function *$p(x)$ if*

$$\int_a^b p(x)\varphi_n(x)\varphi_m(x)\,dx=0 \quad \textit{for } n\neq m \tag{8}$$

and

$$\int_a^b p(x)\varphi_n^2(x)\,dx>0 \quad \textit{for } n=1, 2, \ldots \tag{9}$$

In the case $p(x)\equiv 1$ the definition of orthogonality with weight function turns into the definition of (ordinary) orthogonality given in § 3, Sec. 1 of Chapter 11.

Let $f(x)$ be a function square-integrable on $[a, b]$ with weight function $p(x)$ and let system (3) be orthogonal on $[a, b]$ with weight function $p(x)$.

A series of the form

$$c_1\varphi_1(x)+c_2\varphi_2(x)+\ldots+c_n\varphi_n(x)+\ldots \tag{A}$$

with coefficients c_n, $n=1, 2, \ldots$, determined by the formulas

$$c_n=\frac{\int_a^b p(x)f(x)\varphi_n(x)\,dx}{\int_a^b p(x)\varphi_n^2(x)\,dx} \tag{B}$$

is referred to as the ***Fourier series of the function*** *$f(x)$* ***with respect to system*** *(3), and we write*

$$f(x)\sim c_1\varphi_1(x)+\ldots+c_n\varphi_n(x)+\ldots$$

We say that series (A) ***is convergent in the mean on*** $[a, b]$ ***with weight function*** *$p(x)$* ***to the function*** *$f(x)$ if*

$$\lim_{m\to+\infty}\int_a^b p(x)\Big[f(x)-\sum_{k=1}^m c_k\varphi_k(x)\Big]^2 dx=0 \tag{C}$$

If series (A) converges uniformly or in the mean (with weight function $p(x)$) to the function $f(x)$ on the interval $[a, b]$ its coefficients are uniquely specified by formulas (B). In fact, under the given conditions, the Cauchy-Bunyakovsky inequality implies that

$$\left|\int_a^b p(x)\varphi_n(x)\Big[f(x)-\sum_{k=1}^m c_k\varphi_k(x)\Big]dx\right|\leqslant\left(\int_a^b p(x)\varphi_n^2(x)\,dx\right)^{1/2}\times$$
$$\times\left(\int_a^b p(x)\Big[f(x)-\sum_{k=1}^m c_k\varphi_k(x)\Big]^2 dx\right)^{1/2}\to 0$$

for $m \to +\infty$ and any fixed n. On the other hand, the functions $\varphi_i(x)$, $i=1, 2, \ldots$, being orthogonal on the interval $[a, b]$ with weight function $p(x)$, we have the relation

$$\int_a^b p(x)\varphi_n(x)\left[f(x)-\sum_{k=1}^{m} c_k\varphi_k(x)\right]dx=\int_a^b p(x)f(x)\varphi_n(x)\,dx-$$

$$-c_n\int_a^b p(x)\varphi_n^2(x)\,dx=\text{const}$$

where $m \geqslant n$ and n is fixed. Consequently,

$$\int_a^b p(x)f(x)\varphi_n(x)\,dx-c_n\int_a^b p(x)\varphi_n^2(x)\,dx=0$$

which implies (B).

The definitions of complete and closed systems (see § 6 of Chapter 11) and also the basic theorems related to these notions (see Theorems 11.4-11.7 in §§ 5, 6 of Chapter 11) are easily generalized to the case of systems orthogonal with weight function.

In mathematical physics expansions of functions into series with respect to systems of functions orthogonal with a weight function are widely applied to various problems. Among the most important systems of functions orthogonal with a weight function we can mention various systems of special polynomials (which will be discussed at the end of this appendix) and also the systems of eigenfunctions used in studying the problems of vibration of a circular or a ring-shaped membrane, the systems of eigenfunctions for a sphere and for a spherical layer etc. (see [17]).

Systems of functions orthogonal with weight function (and, in particular, with weight function $p(x)=1$) are especially convenient for expanding functions into series because the corresponding coefficients of such expansions are easily found.

An orthogonal (with weight function) system of functions can be constructed by applying the so-called *orthogonalization process* to a given system of linearly independent functions.

We say that *functions** $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_n(x)$ *are* ***linearly dependent on*** $[a, b]$ *if there are constants* C_1, C_2, ..., C_n, *not all zero, such that the linear combination of the functions* $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_n(x)$ *with the coefficients* C_1, C_2, ..., C_n *is identically equal to zero on* $[a, b]$ *except possibly at the points of discontinuity of the*

* We remind the reader that all the functions under consideration are supposed to be continuous everywhere on $[a, b]$ except possibly at a finite number of points.

functions $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$:

$$C_1\varphi_1(x)+C_2\varphi_2(x)+\ldots+C_n\varphi_n(x)\equiv 0 \tag{10}$$

If identity (10) *(understood in the above sense) implies that all the coefficients* $C_1, C_2, \ldots, C_n$ *of this linear combination are equal to zero the functions* $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$ *are said to be* ***linearly independent*** *on* $[a, b]$.

If a system of functions $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$ *is orthogonal on the interval* $[a, b]$ *with weight function* $p(x)$ *the functions of this system are* ***linearly independent*** *on* $[a, b]$. Indeed, suppose that

$$C_1\varphi_1(x)+C_2\varphi_2(x)+\ldots+C_n\varphi_n(x)\equiv 0 \quad \text{on} \quad [a, b] \tag{11}$$

everywhere except possibly at a finite number of points. Multiplying equality (11) by $p(x)\varphi_m(x)$, $m=1, 2, \ldots$, integrating with respect to x from a to b and taking into account that the functions $\varphi_1(x), \ldots, \varphi_n(x)$ are orthogonal on $[a, b]$ with weight function $p(x)$, we obtain

$$C_m\int_a^b p(x)\varphi_m^2(x)\,dx=0, \quad m=1, 2, \ldots \tag{12}$$

The integral in (12) being different from zero, we see that $C_m=0$, $m=1, 2, \ldots$. Thus we conclude that $C_1=C_2=\ldots=C_n=0$.

If we are given a system of linearly independent functions $\psi_1(x)$, $\psi_2(x), \ldots, \psi_n(x)$ we can easily construct an orthogonal system (with an arbitrary weight function $p(x)$) $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$ such that

$$\left.\begin{array}{l}
\varphi_1(x)=\psi_1(x)\\
\varphi_2(x)=\psi_2(x)+\lambda_{21}\psi_1(x)\\
\varphi_3(x)=\psi_3(x)+\lambda_{31}\psi_1(x)+\lambda_{32}\psi_2(x)\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\varphi_n(x)=\psi_n(x)+\lambda_{n1}\psi_1(x)+\lambda_{n2}\psi_2(x)+\ldots+\lambda_{n\,n-1}\psi_{n-1}(x)
\end{array}\right\} \tag{13}$$

Before proceeding to prove this assertion we note that relations (13) (provided they are valid) imply that every function $\varphi_k(x)$, $k=1, 2, \ldots, n$ is a linear combination of the functions $\psi_1(x), \ldots, \psi_k(x)$ in which the coefficient in $\psi_k(x)$ is equal to unity and, consequently, $\varphi_k(x)$, $k=1, 2, \ldots, n$, may vanish only at the points of discontinuity of the functions $\psi_1(x), \ldots, \psi_k(x)$. In fact, if otherwise, the functions $\psi_1(x), \ldots, \psi_k(x)$ would be linearly dependent on $[a, b]$, which contradicts the hypothesis. Thus, every function $\varphi_k(x)$, $k=1, 2, \ldots, n$, is different from zero at each of its points of continuity, and hence

$$\int_a^b p(x)\varphi_k^2(x)\,dx>0, \quad k=1, 2, \ldots, n \tag{14}$$

It should also be noted that it follows from relations (13) that every function $\psi_k(x)$, $k = 1, 2, \ldots, n$, is a linear combination of the functions $\varphi_1(x)$, $\varphi_2(x)$, $\ldots$, $\varphi_k(x)$ with the coefficient in $\varphi_k(x)$ equal to unity.

To justify formulas (13) (i.e. to prove the possibility of constructing the desired system $\{\varphi_k(x)\}$ orthogonal on $[a, b]$ with weight function $p(x)$) we must show that the numbers λ_{ij} entering into (13) can in fact be found. This can easily be shown by induction. Moreover, it turns out that formulas (13) uniquely specify the quantities λ_{ij}. Indeed, multiplying both sides of the second equality (13) by $p(x)\varphi_1(x)$ and integrating with respect to x from a to b we derive

$$\int_a^b p(x)\varphi_1(x)\psi_2(x)\,dx = \int_a^b p(x)\varphi_1(x)\varphi_2(x)\,dx +$$

$$+\lambda_{21}\int_a^b p(x)\varphi_1^2(x)\,dx = 0$$

Consequently, by the hypothesis that the functions $\varphi_i(x)$, $i = 1, 2, \ldots, n$, are orthogonal on $[a, b]$ with weight function $p(x)$, we obtain

$$\lambda_{21} = -\frac{\int_a^b p(x)\varphi_1(x)\varphi_2(x)\,dx}{\int_a^b p(x)\varphi_1^2(x)\,dx}$$

and thus the coefficient λ_{21} has been determined (and is uniquely specified by (13)). Now suppose that the coefficients λ_{ij}, $i = 1, 2, \ldots, k-1$, $j = 1, 2, \ldots, i-1$, have already been computed and that the functions $\psi_1(x)$, $\ldots$, $\psi_{k-1}(x)$ thus obtained are pairwise orthogonal on $[a, b]$ with weight function $p(x)$. Then the conditions of orthogonality

$$\int_a^b p(x)\psi_k(x)\varphi_j(x)\,dx = \int_a^b p(x)\varphi_k(x)\varphi_j(x)\,dx +$$

$$+\sum_{i=1}^{k-1}\lambda_{ki}\int_a^b p(x)\varphi_i(x)\varphi_j(x)\,dx =$$

$$= \int_a^b p(x)\varphi_k(x)\varphi_j(x)\,dx + \lambda_{kj}\int_a^b p(x)\varphi_j^2(x)\,dx = 0,$$

$$j = 1, 2, \ldots, k-1$$

yield

$$\lambda_{kj} = -\frac{\int_a^b p(x)\,\psi_k(x)\,\varphi_j(x)\,dx}{\int_a^b p(x)\,\varphi_j^2(x)\,dx}, \quad j = 1, 2, \ldots, k-1$$

This determines the function $\varphi_k(x) = \psi_k(x) + \lambda_{k1}\psi_1(x) + \ldots + \lambda_{kk-1}\psi_{k-1}(x)$, and the system of functions $\varphi_1(x), \ldots, \varphi_k(x)$ is orthogonal on $[a, b]$ with weight function $p(x)$. Thus, the assertion has been proved.

The above process of constructing an orthogonal system $\varphi_1(x), \ldots, \varphi_n(x)$ from a given system $\psi_1(x), \ldots, \psi_n(x)$ of linearly independent functions performed according to formulas (13) is referred to as the **orthogonalization process.** In just the same way this process can be applied to infinite systems of functions.

Let us consider the system of functions

$$1, x, x^2, \ldots, x^n, \ldots \tag{15}$$

consisting of integral nonnegative powers of the argument x. As is known, these functions are linearly independent. Applying to (15) the orthogonalization process for the interval $[-1, 1]$ with the weight function $p(x) \equiv 1$ we obtain a system of polynomials which are orthogonal on $[-1, 1]$. We do not write down the formulas for these polynomials because they differ only in constant factors from Legendre's polynomials determined by the formulas

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^n], \quad n = 1, 2, \ldots; \quad P_0(x) = 1 \tag{16}$$

Relations (16) are known as **Rodrigues'* formulas.**

Taking the weight function $p(x) = \frac{1}{\sqrt{1-x^2}}$ and applying the orthogonalization process to the same system of powers (15) on the interval $[-1, 1]$ we arrive at the system of **Chebyshev's polynomials of the first kind.** If we take the weight function $p(x) = \sqrt{1-x^2}$ we obtain **Chebyshev's polynomials of the second kind.**

Considering system (15) on the semi-infinite interval $[0, +\infty)$ and performing the orthogonalization process with the weight function $p(x) = e^{-x}$ we obtain the system of the so-called **Chebyshev-Laguerre** polynomials.** If we take the weight function $p(x) = x^s e^{-x}$, $s > -1$, the orthogonalization process (on the same interval $[0, +\infty)$) results in the generalized **Chebyshev-Laguerre polynomials.**

---

* Rodrigues, Olinde (1794-1851), a French mathematician.

** Laguerre, Edmond Nicolas (1834-1886), a French mathematician.

Finally, taking, for the same system (15), the interval $-\infty < < x < +\infty$ and the weight function $p(x) = e^{-x^2}$ we obtain the system of the **Chebyshev-Hermite*** **polynomials.**

There are convenient general formulas (similar to Rodrigues' formulas (16)) for all special polynomials.

The systems of orthogonal polynomials enumerated above are widely applied to various problems of mathematical physics (e.g. see [17]).

APPENDIX 3 TO CHAPTER 11

FUNCTIONAL SPACE AND GEOMETRIC ANALOGY

The set of functions $Q[a, b]$ defined in § 6 can be regarded as a **functional space** whose elements are functions. Two functions $\varphi(x)$ and $\psi(x)$ belonging to $Q[a, b]$ are considered as representing the same element ("vector") of this space if they differ at no more than a finite number of points of the interval $[a, b]$. In what follows we shall denote the elements of the space $Q[a, b]$ corresponding to the functions $\varphi(x)$, $\psi(x)$, $\eta(x)$, . . . by φ, ψ, η, . . . omitting the argument x.

The sum $\varphi + \psi$ of two elements φ and ψ and the product $\lambda\varphi$ of an element φ by a number λ are defined as the elements represented by the sum $\varphi(x) + \psi(x)$ and by the product $\lambda\varphi(x)$. Then the space $Q[a, b]$ with the operations of addition and multiplication by a number is analogous to the Euclidean space of all three-dimensional vectors with ordinary operations of addition of vectors and multiplication of vectors by scalars. The zero element 0 of the space $Q[a, b]$ is represented by any function which is identically equal to zero on the interval $[a, b]$ except possibly at a finite number of points.

We now define the **scalar product of two elements** φ **and** ψ belonging to $Q[a, b]$ by putting

$$(\varphi, \psi) = \int_a^b \varphi(x)\psi(x)\,dx \qquad (1$$

One can easily verify that the scalar product thus defined satisfies the ordinary conditions for the scalar product of geometric vectors, namely:

(1) $(\varphi, \psi) = (\psi, \varphi)$,

(2) $(\lambda\varphi, \psi) = \lambda(\varphi, \psi)$ where λ is an arbitrary real number,

(3) $(\varphi, \psi_1 + \psi_2) = (\varphi, \psi_1) + (\varphi, \psi_2)$,

(4) $(\varphi . \varphi) \geqslant 0$, and $(\varphi, \varphi) = 0$ if and only if $\varphi = 0$. When pr

---

* Hermite, Charles (1822-1901), a French mathematician.

ving Theorem 11.6 in § 6, Sec. 3, we established the validity of condition (4), and conditions (1)-(3) are evident.

Thus, we see that the space $Q[a, b]$ with the scalar product defined by equality (1) is closely analogous to the Euclidean space of all three-dimensional vectors whose scalar product is defined in the ordinary way.

The "vectors" φ and ψ of $Q[a, b]$ are said to be **orthogonal** if their scalar product is equal to zero, that is

$$(\varphi, \psi) = \int_a^b \varphi(x)\psi(x)\,dx = 0 \tag{2}$$

The norm or "length" of a vector $\varphi \in Q[a, b]$ can now be defined by the equality

$$\|\varphi\| = \sqrt{(\varphi, \varphi)} \tag{3}$$

If $\|\varphi\| \neq 0$ we can put $\psi(x) = \frac{\varphi(x)}{\|\varphi\|}$ and thus obtain, according to Definition 3, the relation

$$\|\psi\|^2 = \left(\frac{\varphi}{\|\varphi\|}, \frac{\varphi}{\|\varphi\|}\right) = \frac{1}{\|\varphi\|^2}(\varphi, \varphi) = 1 \tag{4}$$

which means that ψ is a "unit" vector. The cosine of the angle between $f(x)$ and $g(x)$ is defined by the relation

$$\cos(\widehat{f, g}) = \frac{(f, g)}{\|f\|\,\|g\|} \tag{5}$$

The justification of this definition lies in the fact that, according to the Cauchy-Bunyakovsky inequality (see § 6, Sec. 2 of Chapter 8), we always have

$$|(f, g)| \leqslant \|f\|\,\|g\| \tag{6}$$

and hence relation (5) actually determines a unique angle in the interval $[0, \pi]$.

The **projection** of f on g (where $g \neq 0$) is defined as the scalar quantity

$$\|f\|\cos(\widehat{f, g}) = \frac{(f, g)}{\|g\|} \tag{7}$$

Now let us define the notion of *convergence* for the space $Q[a, b]$. We say that φ_n converges to φ, i.e. $\varphi_n \to \varphi$ for $n \to +\infty$, if

$$\|\varphi_n - \varphi\| = \left(\int_a^b |\varphi_n(x) - \varphi(x)|^2\,dx\right)^{1/2} \to 0 \tag{8}$$

for $n \to -\infty$. Hence, the relation $\lim\limits_{n \to +\infty} \varphi_n = \varphi$ means that the functional sequence $\{\varphi_n(x)\}$ *converges in the mean to* $\varphi(x)$ *on* $[a, b]$.

Similarly, the relation

$$f \doteq f_1 + f_2 + \dots + f_n + \dots \tag{9}$$

is understood in the sense that

$$\left\| f - \sum_{k=1}^{n} f_k \right\| = \left(\int_a^b \left[f(x) - \sum_{k=1}^{n} f_k(x) \right]^2 dx \right)^{1/2} \to 0 \tag{10}$$

for $n \to +\infty$, that is the series $\sum_{k=1}^{+\infty} f_k(x)$ is *convergent in the mean to $f(x)$ on $[a, b]$*.

We now proceed to establish the analogy between the resolution of a vector $\mathbf{x}$ of the three-dimensional Euclidean space with respect to an orthogonal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and the expansion of a function $f(x) \in Q[a, b]$ into a Fourier series with respect to a complete orthogonal system

$$\varphi_1(x),\ \varphi_2(x),\ \dots,\ \varphi_n(x),\ \dots;\quad \varphi_k(x) \in Q[a, b],$$
$$k = 1, 2, \dots \tag{11}$$

For every vector $\mathbf{x}$ belonging to the three-dimensional Euclidean space there exists a unique resolution of the form

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \tag{12}$$

with respect to any fixed orthogonal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. The coefficients of this resolution can be easily found if we use the notion of scalar product. Indeed, multiplying scalarly equality (12) by $\mathbf{e}_i$, $i = 1, 2, 3$, we obtain

$$(\mathbf{x}, \mathbf{e}_i) = x_i(\mathbf{e}_i, \mathbf{e}_i) = x_i \|\mathbf{e}_i\|^2, \quad i = 1, 2, 3 \tag{13}$$

since the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is supposed to be orthogonal.* From (13) we derive

$$x_i = \frac{(\mathbf{x}, \mathbf{e}_i)}{\|\mathbf{e}_i\|^2}, \quad i = 1, 2, 3 \tag{14}$$

The quantities

$$x_k \|\mathbf{e}_k\| = \frac{(\mathbf{x}, \mathbf{e}_k)}{\|\mathbf{e}_k\|}, \quad k = 1, 2, 3 \tag{15}$$

(where the symbol $\|\ \ \|$ designates the length of a vector) are the projections of the vector $\mathbf{x}$ on the axes whose directions are specified by the vectors $\mathbf{e}_k$, $k = 1, 2, 3$.

Taking the scalar squares of both sides of equality (12) we arrive at the relation

$$\|\mathbf{x}\|^2 = x_1^2\|\mathbf{e}_1\|^2 + x_2^2\|\mathbf{e}_2\|^2 + x_3^2\|\mathbf{e}_3\|^2 \tag{16}$$

---

* In the general case the orthogonal vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ are of arbitrary lengths.

which expresses the *Pythagoras** *theorem* for the three-dimensional case: the square of the length of a vector is equal to the sum of the squares of its projections on any three mutually orthogonal axes.

For the space $Q\,[a, b]$ we have a completely analogous situation: every "vector" f can be uniquely expanded with respect to the "vectors" of a complete orthogonal system $\{\varphi_i\}$, i.e. represented in the form

$$f \doteq c_1\varphi_1 + c_2\varphi_2 + \ldots + c_k\varphi_k + \ldots \tag{17}$$

(see § 6, Sec. 1 of Chapter 11) where the coefficients c_k of expansion (17) are determined by the formulas

$$c_k = \frac{(f, \varphi_k)}{\|\varphi_k\|}, \quad k = 1, 2, 3, \ldots \tag{18}$$

Thus, every "vector" f belonging to $Q\,[a, b]$ is uniquely specified by the infinite sequence of its "coordinates" $\{c_k\}$ in every "orthogonal basis" $\{\varphi_k\}$.

The system $\{\varphi_n\}$ being complete, we have *Parseval's relation*

$$\|f\|^2 = c_1^2\,\|\varphi_1\|^2 + c_2^2\,\|\varphi_2\|^2 + \ldots + c_n^2\,\|\varphi_n\|^2 + \ldots \tag{19}$$

(see § 6, Sec. 2) which can be interpreted as **Pythagoras' theorem for the functional space** $Q\,[a, b]$.

If we take not all the base vectors $\mathbf{e}_i$, $i = 1, 2, 3$, but only some of them, for instance, $\mathbf{e}_1$ and $\mathbf{e}_2$, equality (16) is replaced by the inequality

$$\|\mathbf{x}\|^2 \geqslant x_1^2\|\mathbf{e}_1\|^2 + x_2^2\,\|\mathbf{e}_2\|^2 \tag{20}$$

Similarly, if an orthogonal system $\{\varphi_n\}$ of the space $Q\,[a, b]$ is not complete, Parseval's relation (19) is replaced by Bessel's inequality

$$\|f\|^2 \geqslant c_1^2\,\|\varphi_1\|^2 + c_2^2\,\|\varphi_2\|^2 + \ldots + c_n^2\,\|\varphi_n\|^2 + \ldots \tag{21}$$

These geometric ideas and analogues are used in the theory of the so-called **Hilbert**** **spaces** which is widely applied to various problems of quantum mechanics and mathematical physics.

APPENDIX 4 TO CHAPTER 11

SOME APPLICATIONS OF FOURIER TRANSFORMS

We now discuss some applications of Fourier's transformation.

Many physical devices can be interpreted as the so-called **input-output systems**. An *input signal* described by functions $f_1(t)$, $f_2(t)$, ... is applied to the input of such a system which results in the appearan-

---

* Pythagoras, a Greek philosopher and mathematician of the 6th century B.C.

** Hilbert, David (1862-1943), a famous German mathematician.

ce of an *output signal* described by another system of functions $x_1(t), x_2(t), \ldots$. Thus, such a device can be regarded as a convertor which transforms input functions into output functions. For instance, various amplifiers can be regarded as such systems transforming the voltage $f(t)$ of an alternating electric current applied to the input terminals into an alternating voltage appearing at the output terminals.

An input-output system is called **linear** if the following two conditions are fulfilled:

(1) if $f(t)$ is transformed into $x(t)$ then $c\,f(t)$ (where c is an arbitrary constant) is transformed into $c\,x(t)$.

(2) if $f_1(t)$ and $f_2(t)$ are transformed, respectively, into $x_1(t)$ and $x_2(t)$ then $f_1(t)+f_2(t)$ is transformed into $x_1(t)+x_2(t)$.

If conditions (1) and (2) are satisfied we say that the **principle of superposition** holds for the system in question.

We shall also suppose that every steady-state process of harmonic vibration with frequency ω is transformed into another steady-state vibrational process with the same frequency. This means that we impose one more condition:

(3) every function of the form $e^{i\omega t}$ goes into a function of the form $A(\omega)e^{i\omega t}$.

In the general case the factor of proportionality $A(\omega)$ is dependent on the frequency ω, which means that an input-output system may transform harmonic oscillations with different frequencies in a different way. The function $A = A(\omega)$ is called the **spectral characteristic** of the system. In the general case this function may assume complex values:

$$A(\omega) = R(\omega)\,e^{i\varphi(\omega)} \quad \text{where} \quad R(\omega) = |A(\omega)| \quad \text{and}$$

$$\varphi(\omega) = \arg A(\omega)$$

Consequently, harmonic oscillations described by a function $e^{i\omega t}$ are transformed into harmonic oscillations of the form $A(\omega)\,e^{i\omega t} = R(\omega)\,e^{i(\omega t+\varphi(\omega))}$.

The modulus $R(\omega) = |A(\omega)|$ of the spectral characteristic is known as the ***frequency characteristic*** of the system. The amplitude of the output harmonic signal is $R(\omega)$ times that of the input harmonic signal with the given frequency ω. The argument $\varphi(\omega) = \arg A(\omega)$ of the spectral characteristic $A(\omega)$ is called the ***phase characteristic*** of the input-output system. It determines the phase displacement of a harmonic signal of a given frequency ω when it is transformed by the system.

If the spectral characteristic of a linear input-output system is known we can solve the following two problems:

1. Given an input function $f(t)$, it is required to determine the corresponding output function $x(t)$.

2. It is necessary to find the input function $f(t)$ from a given output function $x(t)$.

The second problem is reverse with respect to the first. We shall begin with the first problem. Let an input signal described by a function $f(t)$ be applied to the input of the system. The Fourier transform of the function $f(t)$ is determined by the relation

$$\overline{f}(\omega)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(\tau)\, e^{-i\omega\tau}\, d\tau \tag{1}$$

and the function $f(t)$, the inverse transform, is represented by Fourier's inversion formula:

$$f(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \overline{f}(\omega)\, e^{i\omega t}\, d\omega \tag{2}$$

The integral on the right-hand side of equality (2) can be regarded as a sum of infinitely many infinitesimal harmonic oscillations of the form

$$\frac{1}{\sqrt{2\pi}}\overline{f}(\omega)\, e^{i\omega t}\, d\omega \tag{3}$$

But every function of the form $e^{i\omega t}$ is transformed into the function $A(\omega)e^{i\omega t}$ and, consequently, according to property 1 of the system the harmonic signal $\left(\frac{1}{\sqrt{2\pi}}\overline{f}(\omega)\, d\omega\right) e^{i\omega t}$ goes into the harmonic signal

$$\left(\frac{1}{\sqrt{2\pi}}\overline{f}(\omega)\, d\omega\right) A(\omega)\, e^{i\omega t}=\frac{1}{\sqrt{2\pi}} A(\omega)\,\overline{f}(\omega)\, e^{i\omega t}\, d\omega \tag{4}$$

Property (2) of the system implies that the sum of harmonic vibrations of form (3) is transformed into the sum of functions (4) and hence the function $f(t)$ determined by relation (2) is transformed into the function

$$x(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} A(\omega)\,\overline{f}(\omega)\, e^{i\omega t}\, d\omega \tag{5}$$

Formula (5) gives the solution of Problem 1.

Relation (5) indicates that the Fourier transform $x(\omega)$ of the function $x(t)$ is expressed as

$$\overline{x}(\omega)=A(\omega)\,\overline{f}(\omega) \tag{6}$$

Thus, to find the Fourier transform $\overline{x}(\omega)$ of the output function $x(t)$ we must multiply the Fourier transform $\overline{f}(\omega)$ of the input function $f(t)$ by the spectral characteristic of the input-output system.

To solve the reverse Problem 2 we find the function

$$\overline{f}(\omega) = \frac{\overline{x}(\omega)}{A(\omega)} \tag{7}$$

from relation (6) and then apply the Fourier inverse transformation to equality (7), which results in

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\overline{x}(\omega)}{A(\omega)} e^{i\omega t} d\omega \tag{8}$$

Relation (8) expresses the solution of Problem 2. It enables us to reconstruct the input function $f(t)$ corresponding to the output function $x(t)$. To this end, we first apply the Fourier transformation and determine the Fourier transform $\overline{x}(\omega)$ of the function $x(t)$ and then use formula (8).

Such problems are encountered in radiophysics, radio engineering, in the theory of control systems and the like.

Fourier's transformation is also widely applied to solving various boundary value problems of mathematical physics. The matter is, that the Fourier transform of a sought-for function may satisfy an equation which is much simpler than the original equation for the unknown function. Therefore the boundary value problems of mathematical physics are solved by means of the Fourier transformation according to the following scheme: we first apply the Fourier transformation to the equation which is satisfied by the sought-for function and thus obtain an equation for the Fourier transform of the unknown function, then we find the Fourier transform of the sought-for function from the equation thus obtained and finally determine the solution of the original problem by means of the Fourier inverse transformation (see [17]).

Let us consider an example. Suppose that it is required to find the distribution of temperature in an infinite rod at an arbitrary moment of time $t > 0$ from a given temperature distribution at the moment $t = 0$. Let the x-axis be directed along the rod. Then the temperature u at a point x of the rod at moment t is described by a function $u = u(x, t)$, $-\infty < x < +\infty$, $0 < t < +\infty$, of two variables. As is known (e.g. see [17]), the temperature $u(x, t)$ satisfies the **heat conductivity equation**

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < +\infty, \quad t > 0 \tag{9}$$

The initial temperature distribution being known, we can write

$$u(x, 0) = f(x), \quad -\infty < x < +\infty \tag{10}$$

where $f(x)$ is a given function. Thus, to determine the distribution of temperature in the unbounded rod at an arbitrary time moment t

we must find the solution of equation (9) which assumes the initial values (10) at the moment $t = 0$.

Let us solve this problem by applying the Fourier transformation to the function $u(x, t)$ regarded as a function of the argument x for every fixed value of t. Designating the Fourier transform of the function $u(x, t)$ by $\bar{u}(\lambda, t)$ we can write

$$\bar{u}(\lambda, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u(x, t)\, e^{-i\lambda x}\, dx \tag{11}$$

Let us multiply both sides of equation (9) by $\frac{1}{\sqrt{2\pi}} e^{-i\lambda x}$ and integrate the resulting relation with respect to x from $-\infty$ to $+\infty$ under the assumption that the function u and its derivatives tend to zero sufficiently fast as $x \to \pm\infty$. Then, integrating by parts, we obtain

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\partial u}{\partial t} e^{-i\lambda x}\, dx = \frac{\partial}{\partial t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u(x, t)\, e^{-i\lambda x}\, dx =$$

$$= \frac{d\bar{u}(\lambda, t)}{dt} = a^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\partial^2 u}{\partial x^2} e^{-i\lambda x}\, dx =$$

$$= a^2 \frac{1}{\sqrt{2\pi}} \frac{\partial u}{\partial x} e^{-i\lambda x} \Big|_{x=-\infty}^{x=+\infty} +$$

$$+ a^2 \frac{1}{\sqrt{2\pi}} i\lambda u e^{-i\lambda x} \Big|_{x=-\infty}^{x=+\infty} -$$

$$- a^2\lambda^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u e^{-i\lambda x}\, dx = -a^2\lambda^2 \bar{u}(\lambda, t)$$

since the expression $e^{-i\lambda x}$ is bounded* and (by the above hypothesis) the function u and its derivative $\frac{\partial u}{\partial x}$ tend to zero when $x \to \pm\infty$. Consequently, the Fourier transform of the sought-for function satisfies the equation

$$\frac{d\bar{u}}{dt} + a^2\lambda^2 \bar{u} = 0 \tag{12}$$

which is considerably simpler than equation (9). From equality (11) we find, by putting $t = 0$, the initial condition for $\bar{u}(\lambda, t)$:

$$\bar{u}(\lambda, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u(x, 0)\, e^{-i\lambda x}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, e^{-i\lambda x}\, dx = \bar{f}(\lambda) \tag{13}$$

---

* In fact, we have $|e^{-i\lambda x}| = |\cos\lambda x + i \sin\lambda x| = \sqrt{\cos^2\lambda x + \sin^2\lambda x} = 1$.

Let us solve equation (12) with initial condition (13). From (12 we find

$$\frac{d\bar{u}}{\bar{u}} = -a^2\lambda^2\,dt$$

and, consequently, we have

$$\ln \bar{u} = -a^2\lambda^2 t + \ln C$$

and

$$\bar{u}(\lambda, t) = Ce^{-a^2\lambda^2 t}$$

where C is independent of t and may only depend on λ. The quantity C can be determined by means of initial condition (13):

$$\bar{u}(\lambda, 0) = C = \bar{f}(\lambda)$$

Substituting the value of C thus found into the foregoing equalit we obtain the following expression for the Fourier transform of th unknown function:

$$\bar{u}(\lambda, t) = \bar{f}(\lambda)\, e^{-a^2\lambda^2 t} \tag{14}$$

To determine $u(x, t)$, we apply the Fourier inverse transfor mation to equality (14) in which we substitute the expressio $\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} f(\xi)\, e^{-i\lambda\xi}\,d\xi$ for $\bar{f}(\lambda)$. This results in

$$u(x, t) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} \bar{u}(\lambda, t)\, e^{i\lambda x}\,d\lambda =$$

$$= \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} f(\xi)\,d\xi \int\limits_{-\infty}^{+\infty} e^{-a^2\lambda^2 t} e^{i\lambda(x-\xi)}\,d\lambda =$$

$$= \frac{1}{\pi}\int\limits_{-\infty}^{+\infty} f(\xi)\,d\xi \int\limits_{0}^{+\infty} e^{-a^2\lambda^2 t}\cos\lambda(x-\xi)\,d\xi =$$

$$= \frac{1}{2a\sqrt{\pi t}}\int\limits_{-\infty}^{+\infty} f(\xi)\, e^{-\frac{(x-\xi)^2}{4a^2 t}}\,d\xi$$

because we have (see Chapter 10, § 2, Sec. 5)

$$\int\limits_{0}^{+\infty} e^{-\alpha^2\lambda^2}\cos\beta\lambda\,d\lambda = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}\, e^{-\frac{\beta^2}{4\alpha}}$$

Thus, the solution of equation (9) with initial conditions (10) i expressed by the formula

$$u(x, t) = \frac{1}{2a\sqrt{\pi t}}\int\limits_{-\infty}^{+\infty} f(\xi)\, e^{-\frac{(x-\xi)^2}{4a^2 t}}\,d\xi$$

Fourier's sine and cosine transforms are similarly applied to solving various boundary value problems for the semi-infinite interval $0 \leqslant x < +\infty$ (e.g. see [17]).

The application of Fourier's multiple integrals and the corresponding Fourier transforms makes it possible to solve some boundary value problems for unbounded regions in the plane and in space, for instance, such as the whole plane, half-plane, a quadrant, the entire three-dimensional space, half-space etc.

APPENDIX 5 TO CHAPTER II

EXPANDING DELTA FUNCTION IN FOURIER SERIES AND FOURIER INTEGRAL

Let us take the delta function $\delta(x_0, x)$ and compute (formally) its Fourier coefficients by using the ordinary formulas. This yields, for $x_0 \in (-l, l)$, the expressions

$$a_k = \frac{1}{l}\int_{-l}^{l} \delta(x_0, \xi)\cos\frac{k\pi\xi}{l}\,d\xi = \frac{1}{l}\cos\frac{k\pi x_0}{l}, \quad k = 1, 2, \ldots$$

$$a_0 = \frac{1}{l}\int_{-l}^{l} \delta(x_0, \xi)\,d\xi = \frac{1}{l}$$

$$b_k = \frac{1}{l}\int_{-l}^{l} \delta(x_0, \xi)\sin\frac{k\pi\xi}{l}\,d\xi = \frac{1}{l}\sin\frac{k\pi x_0}{l}, \quad k = 1, 2, \ldots$$

Consequently, (formal) Fourier's series of the delta function $\delta(x_0, x)$ is of the form

$$\delta(x_0, x) \sim \frac{1}{2l} + \frac{1}{l}\sum_{k=1}^{+\infty}\left(\cos\frac{k\pi x_0}{l}\cos\frac{k\pi x}{l} + \sin\frac{k\pi x_0}{l}\sin\frac{k\pi x}{l}\right) =$$

$$= \frac{1}{2l} + \frac{1}{l}\sum_{k=1}^{+\infty}\cos\frac{k\pi}{l}(x - x_0) \tag{1}$$

or, in the complex form,

$$\delta(x_0, x) \sim \frac{1}{2l}\sum_{k=-\infty}^{+\infty} e^{ik\frac{\pi}{l}(x - x_0)} \tag{2}$$

Let us consider the sequence of partial sums of this series:

$$\tilde{\delta}_n(x_0, x) = \frac{1}{2l} + \frac{1}{l}\sum_{k=1}^{n}\left(\cos\frac{k\pi x_0}{l}\cos\frac{k\pi x}{l} + \sin\frac{k\pi x_0}{l}\sin\frac{k\pi x}{l}\right), \tag{3}$$

$$n = 1, 2, \ldots$$

The complex form of this sequence is

$$\tilde{\delta}_n(x_0, x) = \frac{1}{2l} \sum_{k=-n}^{n} e^{ik\frac{\pi}{l}(x-x_0)}, \quad n = 1, 2, \ldots \tag{4}$$

If $f(x)$ is an arbitrary piecewise smooth function on an interval $(-l, l)$ we obviously have

$$\lim_{n\to+\infty} \int_{-l}^{l} f(x)\, \tilde{\delta}_n(x_0, x)\, dx = f(x_0), \quad x_0 \in (-l, l) \tag{5}$$

provided that the function $f(x)$ has been redefined at every point of discontinuity x^* according to the relation $f(x^*) = \frac{f(x^*-0)+f(x^*+0)}{2}$. We can therefore say (see the note after relations (7), (8) and (9) in Appendix 2 to Chapter 8 and footnote on page 381) that the *sequence* $\{\tilde{\delta}_n(x_0, x)\}$ *is weakly convergent to the delta function* $\delta(x_0, x)$ (relative to the class of piecewise smooth functions on $(-l, l)$) or, in other words, *series* (1) *is weakly convergent to* $\delta(x_0, x)$. This fact can be expressed symbolically by the relation

$$\delta(x_0, x) = \frac{1}{2l} + \frac{1}{l} \sum_{k=1}^{+\infty} \left(\cos\frac{k\pi x_0}{l} \cos\frac{k\pi x}{l} + \sin\frac{k\pi x_0}{l} \sin\frac{k\pi x}{l} \right) \tag{6}$$

Multiplying both sides of equality (6) by an arbitrary piecewise smooth function $f(x)$ and performing term-by-term integration with respect to x from $-l$ to l we arrive at the equality

$$f(x_0) = \frac{a_0}{2} + \sum_{k=1}^{+\infty} \left(a_k \cos\frac{k\pi x_0}{l} + b_k \sin\frac{k\pi x_0}{l} \right) \tag{7}$$

where the coefficients a_k, b_k are determined by the ordinary formulas for Fourier's coefficients of the trigonometric series of the piecewise smooth function $f(x)$ on the interval $(-l, l)$. The validity of equality (7) was proved in § 2, Sec. 4 of Chapter 11.

Equality (6) understood in the sense of weak convergence is called **the expansion of the delta function $\delta(x_0, x)$ in the Fourier series with respect to the trigonometric system.**

Similarly, the delta function $\delta(x_0, x)$ can be written in the form of Fourier's integral. Applying (formally) the Fourier integral formula (see § 9) to $\delta(x_0, x)$ we obtain

$$\delta(x_0, x) = \frac{1}{\pi} \int_{0}^{+\infty} d\lambda \int_{-\infty}^{+\infty} \delta(x_0, \xi) \cos\lambda(\xi - x)\, d\xi$$

which results in

$$\delta(x_0, x) = \frac{1}{\pi} \int_0^{+\infty} \cos \lambda (x_0 - x)\, d\lambda. \tag{8}$$

This relation is called the **representation of the delta function** $\delta(x_0, x)$ **as a Fourier integral.** It should also be understood in the sense of weak convergence. Multiplying both sides of (8) by an arbitrary function $f(x)$ (absolutely integrable over the x-axis from $-\infty$ to $+\infty$ and piecewise smooth on every finite interval) and reversing the order of integration with respect to x and λ on the right-hand side we arrive at the relation

$$f(x_0) = \frac{1}{\pi} \int_0^{+\infty} d\lambda \int_{-\infty}^{+\infty} f(\xi) \cos \lambda (x_0 - \xi)\, d\xi \tag{9}$$

which was proved, for a function $f(x)$ of this type, in § 9, Sec. 2 of Chapter 11.

Thus, when expanding the delta function $\delta(x_0, x)$ in Fourier series (6) or representing it as Fourier integral (8) we can operate on $\delta(x_0, x)$ as if it were an ordinary piecewise smooth function (absolutely integrable from $-\infty$ to $+\infty$ in the case of Fourier's integral). We can also say that expansions (6) and (8) (understood in the sense of weak convergence) can be dealt with, to some extent, as ordinary equalities.

For greater detail on the delta function we refer the reader to [7], [9] and [17].

APPENDIX 6 TO CHAPTER 11

UNIFORM APPROXIMATION OF FUNCTIONS WITH POLYNOMIALS

Here we present another proof of Weierstrass' polynomial approximation theorem (see § 5 of Chapter 11) on approximating a continuous function with algebraic polynomials which can easily be extended to functions of several independent variables. In the case of a function $f(x)$ possessing the continuous derivatives $f^{(s)}(x)$, $s = 1, \ldots, N$, this proof makes it possible to construct an algebraic polynomial approximating uniformly the function f in such a way that the derivatives of this polynomial up to the order N inclusive uniformly approximate the corresponding derivatives of the function f.

Theorem 1 (Weierstrass' Polynomial Approximation Theorem for Functions of One Independent Variable).

Any function $f(x)$ continuous on a bounded closed interval $a \leqslant x \leqslant b$ can be uniformly approximated with an algebraic polynomial to an arbitrary degree of accuracy.

Proof. We shall suppose, without loss of generality, that the interval $[a, b]$ is such that $0 < a < b < 1$ (if otherwise, we can perform the corresponding change of the independent variable x). Let us take arbitrary numbers α and β such that $0 < \alpha < a < b < < \beta < 1$. Now, considering these α and β to be fixed, let us continuously extend the function $f(x)$ to the whole interval $0 \leqslant x \leqslant 1$ in such a way that $f(x)$ is identically equal to zero ($f(x) \equiv 0$) for $0 \leqslant x \leqslant \alpha$ and $\beta \leqslant x \leqslant 1$.

We shall prove that the expression

$$P_n(x) = \frac{\int\limits_{\alpha}^{\beta} f(u)\,[1-(u-x)^2]^n\,du}{\int\limits_{-1}^{1} (1-u^2)^n\,du} \tag{1}$$

(which is an algebraic polynomial in x of degree $2n$) uniformly approximates, for a sufficiently large n, the function $f(x)$ on the interval $[a, b]$ to an arbitrary accuracy. To this end we note that

$$J_n = \int\limits_0^1 (1-v^2)^n\,dv \geqslant \int\limits_0^1 (1-v)^n\,dv = -\frac{(1-v)^n}{n+1}\Bigg|_{v=0}^{v=1} = \frac{1}{n+1}$$

and

$$J_n^* = \int\limits_{\delta}^{1} (1-v^2)^n\,dv < (1-\delta^2)^n$$

for any δ, $0 < \delta < 1$. Consequently, we have

$$0 < \frac{J_n^*}{J_n} \leqslant (1-\delta^2)^n\,(n+1) \to 0 \quad \text{for } n \to +\infty \tag{2}$$

provided that $\delta = \text{const}$, $0 < \delta < 1$. Performing the change of variable $u - x = v$ we can rewrite expression (1) in the form

$$P_n(x) = \frac{\int\limits_{\alpha-x}^{\beta-x} f(v+x)\,(1-v^2)^n\,dv}{\int\limits_{-1}^{1} (1-v^2)^n\,dv} \tag{3}$$

Let us now estimate the difference

$$P_n(x) - f(x) = \frac{\int\limits_{\alpha-x}^{\beta-x} f(v+x)\,(1-v^2)^n\,dv - \int\limits_{-1}^{1} f(x)\,(1-v^2)^n\,dv}{2J_n} \tag{4}$$

on the interval $a \leqslant x \leqslant b$. Given $\varepsilon > 0$, we can choose δ, $0 < \delta < 1$, such that the inequality

$$|f(x+v)-f(x)| < \frac{\varepsilon}{2} \quad \text{for } a \leqslant x \leqslant b \quad \text{and} \quad |v| \leqslant \delta \tag{5}$$

is fulfilled and the relation $0 < x+v < 1$ holds for $a \leqslant x \leqslant b$ and $|v| \leqslant \delta$. The numerator of fraction (4) can be represented as

$$\int_{\alpha-x}^{-\delta} f(v+x)(1-v^2)^n\,dv + \int_{\delta}^{\beta-x} f(v+x)(1-v^2)^n\,dv$$
$$+\int_{-\delta}^{\delta} [f(x+v)-f(x)](1-v^2)^n\,dv + \int_{-1}^{-\delta} f(x)(1-v^2)^n\,dv +$$
$$+\int_{\delta}^{1} f(x)(1-v^2)^n\,dv \tag{6}$$

Furthermore, by (5) we have

$$\left|\int_{-\delta}^{\delta} [f(x+v)-f(x)](1-v^2)^n\,dv\right| \leqslant 2\,\frac{\varepsilon}{2}\,[J_n - J_n^*] \tag{7}$$

Now putting $M = \max\limits_{a \leqslant x \leqslant b} |f(x)|$ we obtain the estimations

$$\left|\int_{\alpha-x}^{-\delta} f(v+x)(1-v^2)^n\,dv\right| \leqslant MJ_n^*$$

and (8)

$$\left|\int_{\delta}^{\beta-x} f(v+x)(1-v^2)^n\,dv\right| \leqslant MJ_n^*$$

since we have $-1 < \alpha - x < 0$ and $0 < \beta - x < 1$ for $a \leqslant x \leqslant b$ and the inequalities

$$\left|\int_{-1}^{\delta} f(x)(1-v^2)^n\,dv\right| \leqslant MJ_n^* \quad \text{and} \quad \left|\int_{\alpha}^{1} f(x)(1-v^2)^n\,dv\right| \leqslant MJ_n^* \tag{9}$$

Therefore the numerator of fraction (4) does not exceed the quantity

$$2\,\frac{\varepsilon}{2}\,J_n + 4MJ_n^* \tag{10}$$

in its modulus. Hence, for (4) we can write the inequality

$$| P_n(x) - f(x) | < \frac{2\frac{\varepsilon}{2} J_n + 4MJ_n^*}{2J_n} = \frac{\varepsilon}{2} + 2M \frac{J_n^*}{J_n} \tag{11}$$

for all $x \in [a, b]$. But, by (2), the second summand on the right-hand side of inequality (11) is less than $\frac{\varepsilon}{2}$ for all sufficiently large n. Consequently, for all sufficiently large n we have

$$| P_n(x) - f(x) | < \varepsilon \quad \text{for all} \quad x \in [a, b] \tag{12}$$

and hence the theorem has been proved.

For functions of several variables we can similarly prove

Theorem 2 (Weierstrass' Polynomial Approximation Theorem for Functions of Several Independent Variables). *Suppose a function $f(x_1, x_2, \ldots, x_m)$ is continuous in an m-dimensional parallelepiped Π: $a_i \leqslant x_i \leqslant b_i$, $i = 1, 2, \ldots, m$, where $0 < a_i < b_i < 1$. Let us continuously extend f to the entire m-dimensional unit cube E_m: $0 \leqslant x_i \leqslant 1$, $i = 1, 2, \ldots, m$, in such a way that f is identically equal to zero outside a wider m-dimensional parallelepiped Π^*: $\alpha_i \leqslant \leqslant x_i \leqslant \beta_i$ where $0 < \alpha_i < a_i < b_i < \beta_i < 1$, $i = 1, 2, \ldots, m$. Then the algebraic polynomial of degree nm in the variables $x_1, x_2, \ldots \ldots, x_m$ expressed by the formula*

$$P_n(x_1, x_2, \ldots, x_m) =$$

$$= \frac{\int\limits_{\alpha_1}^{\beta_1} \cdots \int\limits_{\alpha_m}^{\beta_m} f(u_1, \ldots, u_m)[1-(u_1-x_1)^2]^n \ldots [1-(u_m-x_m)^2]^n \, du_1 \ldots du_m}{\left[\int\limits_{-1}^{1} (1-u^2)^n \, du\right]^m} \tag{13}$$

uniformly approximates the function $f(x_1, \ldots, x_m)$ to an arbitrarily chosen accuracy in the domain Π for sufficiently large n.

Note 1. If the function $f(x)$ ($f(x_1, \ldots, x_m)$) possesses continuous derivatives up to an order N inclusive the derivatives of $P_n(x)$ ($P_n(x_1, \ldots, x_m)$) up to the order N inclusive uniformly approximate the corresponding derivatives of $f(x)$ ($f(x_1, \ldots, x_m)$) on the interval $[a, b]$ (in the parallelepiped Π) to an arbitrary degree of accuracy for all sufficiently large n.

We shall illustrate the proof of this assertion by taking the simple case of a function $f(x)$ which is continuous on an interval $a \leqslant x \leqslant b$, $0 < a < b < 1$, and has the continuous derivative $f'(x)$ on $[a, b]$. Let us extend $f(x)$ to the whole interval $0 \leqslant x \leqslant 1$ in such a way that $f(x)$ and $f'(x)$ are continuous on this interval and identically equal to zero outside an interval $\alpha \leqslant x \leqslant \beta$ where $0 < \alpha < a < < b < \beta < 1$. Differentiating polynomial (1) with respect to x and

integrating by parts we obtain

$$P'_n(x) = \frac{\dfrac{d}{dx}\displaystyle\int_{\alpha}^{\beta} f(u)\,[1-(u-x)^2]^n\,du}{2J_n} = \frac{\displaystyle\int_{\alpha}^{\beta} f(u)\,\frac{d}{dx}\,[1-(u-x^2]^n\,du}{2J_n} =$$

$$= \frac{-\displaystyle\int_{\alpha}^{\beta} f(u)\,\frac{d}{du}\,[1-(u-x)^2]^n\,du}{2J_n} = \frac{\displaystyle\int_{\alpha}^{\beta} f'(u)\,[1-(u-x)^2]^n\,du}{2J_n} =$$

$$= \frac{\displaystyle\int_{\alpha-x}^{\beta-x} f'(v+x)\,(1-v^2)^n\,dv}{\displaystyle\int_{-1}^{1}(1-v^2)^n\,dv}$$

Now the difference

$$P'_n(x) - f'(x) = \frac{\displaystyle\int_{\alpha-x}^{\beta-x} f'(x+v)\,(1-v^2)^n\,dv - \int_{-1}^{1} f'(x)\,(1-v^2)^n\,dv}{2J_n}$$

can be estimated as was done for difference (4) in Theorem 1, and this completes the proof of our assertion.

Note 2. The theorem below is a direct consequence of Theorem 2.

***Theorem 3** (**Weierstrass' Trigonometric Approximation Theorem**). If $f(x)$ is a continuous function on an interval $l \leqslant x \leqslant l$ which assumes equal values at its end points, i.e. $f(-l) = f(l)$, it can be uniformly approximated on this interval, to an arbitrary accuracy, by a trigonometric polynomial of the form*

$$\frac{a_0}{2} + \sum_{k=1}^{m}\left(a_k \cos\frac{k\pi x}{l} + b_k \sin\frac{k\pi x}{l}\right) \tag{14}$$

Proof. Let us put $\frac{\pi x}{l} = \theta$. Then the function $F(\theta) = f\left(\frac{l\theta}{\pi}\right) = f(x)$ is continuous in the interval $-\pi \leqslant \theta \leqslant \pi$, and $F(-\pi) = F(\pi)$. Let us introduce, on the plane with Cartesian coordinates ξ, η, the polar coordinates θ, ρ: $\xi = \rho\cos\theta$, $\eta = \rho\sin\theta$. Now consider the function $\varphi(\xi, \eta) = \rho F(\theta)$. This function is continuous throughout the ξ, η-plane and coincides with $F(\theta)$ for $\rho = 1$, that is $\varphi(\xi, \eta) = F(\theta)$ on the circumference of the circle $\xi^2 + \eta^2 \leqslant 1$. By Theorem 2, the function $\varphi(\xi, \eta)$ can be uniformly approximated in the square $-1 \leqslant \xi \leqslant 1$, $-1 \leqslant \eta \leqslant 1$ by an algebraic polynomial $P_n(\xi, \eta)$ to an arbitrary accuracy. Hence, putting $\rho = 1$ we arrive

at the trigonometric polynomial P_n (cos θ, sin θ) which approximates the function $f(x)$ with the same accuracy on the interval $-\pi \leqslant \leqslant \theta \leqslant \pi$. Next we return to the argument $x = \frac{l\theta}{\pi}$ and obtain the trigonometric polynomial $P_n\left(\cos\frac{\pi x}{l}, \sin\frac{\pi x}{l}\right)$ which uniformly approximates the function $f(x)$ on the interval $-l \leqslant x \leqslant l$ with the same accuracy. Replacing in the latter polynomial the products and the higher powers of $\cos\frac{\pi x}{l}$ and $\sin\frac{\pi x}{l}$ by the corresponding linear combinations of sines and cosines of multiple arcs we reduce this polynomial to form (14). The theorem has been proved.

APPENDIX 7 TO CHAPTER 11

ON STABLE SUMMATION OF FOURIER SERIES WITH PERTURBED COEFFICIENTS

Suppose that we know the exact values c_k of the Fourier coefficients of a square-integrable function $f(x)$ on an interval $[a, b]$ with respect to an orthonormal* system $\{\varphi_k(x)\}$:

$$c_k = \int_a^b f(x)\,\varphi_k(x)\,dx, \quad k = 1, 2, \ldots \tag{1}$$

If the equality

$$f(x) = \sum_{k=1}^{+\infty} c_k\varphi_k(x) \tag{2}$$

is fulfilled for a value of x belonging to $[a, b]$ we can replace in (2) the sum of the series by its nth partial sum and thus obtain the approximate equality

$$f(x) \approx \sum_{k=1}^{n} c_k\varphi_k(x) \tag{3}$$

which goes into exact equality (2) when $n \to +\infty$. Thus, increasing the number of terms n we can obtain equality (3) which approximates exact equality (2) at the point x to an arbitrary accuracy.

But in practical problems we usually deal with approximate values of the Fourier coefficients, i.e. with the numbers

$$\tilde{c}_k = c_k + \Delta c_k \tag{4}$$

* An orthogonal system $\{\varphi_k(x)\}$ is called orthonormal if the norms of all the functions $\varphi_k(x)$, $k = 1, 2, \ldots$, are equal to unity, i.e. $\|\varphi_k\| = \int_a^b \varphi_k^2(x)\,dx = 1$, $k = 1, 2, \ldots$ — *Tr.*

Quantities (4) are termed *perturbed coefficients* and the discrepancies Δc_k are the *errors* or *perturbations*.

If we substitute the approximate values $\widetilde{c}_k$ for the exact values c_k of the Fourier coefficients into approximate equality (3) this will result in the approximate equality

$$f(x) \approx \sum_{k=1}^{n} \widetilde{c}_k \varphi_k(x) \tag{5}$$

whose accuracy may even be worsened when the number of terms n is too great. If the perturbations Δc_k are not subjected to any restrictions this assertion appears quite clear. Usually we impose the following restriction on the quantities Δc_k:

$$\sum_{k=1}^{+\infty} (\Delta c_k)^2 < \delta^2 \tag{6}$$

where δ^2 is a sufficiently small number. If condition (6) is fulfilled for a given finite value of δ it follows that there exists a square integrable function $\widetilde{f}(x)$ on $[a, b]$ for which the series $\sum_{k=1}^{+\infty} \widetilde{c}_k \varphi_k(x)$ is its Fourier series* and, besides, by Parseval's relation

$$\int_a^b [\widetilde{f}(x) - f(x)]^2 \, dx = \sum_{k=1}^{+\infty} (\widetilde{c}_k - c_k)^2 \quad \sum_{k=1}^{+\infty} (\Delta c_k)^2 < \delta^2 \tag{7}$$

the mean square deviation of $\widetilde{f}(x)$ from $f(x)$ on $[a, b]$ is less than δ^2.

But the validity of condition (6) for an arbitrarily small δ does not guarantee the convergence of series

$$\sum_{k=1}^{+\infty} \widetilde{c}_k \varphi_k(x) \tag{8}$$

---

* Indeed, inequality $\sum_{k=1}^{+\infty} (\Delta c_k)^2 < +\infty$, Bessel's inequality $\sum_{k=1}^{+\infty} c_k^2 < \int_a^b f^2(x)\, dx$ and the evident inequality $\widetilde{c}_k^2 \leqslant 2[c_k^2 + (\Delta c_k)^2]$, $k = 1, 2, \ldots$, imply the convergence of the series $\sum_{k=1}^{+\infty} \widetilde{c}_k^2$. But in the theory of functions of a real argument (e.g. see the Riesz-Fisher theorem in [11]) it is proved that a series of form $\sum_{k=1}^{+\infty} \widetilde{c}_k^2$ is convergent if and only if there exists a function $\widetilde{f}(x)$ square-integrable on $[a, b]$ for which the series $\sum_{k=1}^{+\infty} \widetilde{c}_k \varphi_k(x)$ is its Fourier series.

with the perturbed coefficients. Therefore, even if condition (6) is fulfilled, the accuracy of approximate equality (5) may even be worsened if the number of terms n is too large.

For example, let us consider the complete orthonormal system

$$\varphi_1(x)=\frac{1}{\sqrt{\pi}},\quad \varphi_{k+1}(x)=\sqrt{\frac{2}{\pi}}\cos kx,\quad k=1,\ 2,\ \ldots \tag{9}$$

on the interval $0\leqslant x\leqslant \pi$. Suppose that a function $f(x)$ is arbitrarily smooth on the interval $[0\leqslant x\leqslant \pi]$ and that its exact Fourier series

$$\sum_{k=1}^{+\infty} c_k\varphi_k(x)=\frac{c_1}{\sqrt{\pi}}+\sum_{k=1}^{+\infty} c_{k+1}\sqrt{\frac{2}{\pi}}\cos kx \tag{10}$$

converges to it arbitrarily fast at the point $x=0$. Now let us put

$$\Delta c_1=0,\quad \Delta c_{k+1}=\frac{\delta'}{k}\quad \text{for } k=1,\ 2,\ \ldots,\quad \frac{\pi^2\delta'^2}{6}<\delta^2,\quad \delta'>0 \tag{11}$$

Then we have

$$\sum_{k=1}^{+\infty}(\Delta c_k)^2=\frac{\pi^2\delta'^2}{6}<\delta^2 \tag{12}$$

which means that condition (6) is fulfilled for the perturbations Δc_k. But the series

$$\sum_{k=1}^{+\infty}\tilde{c}_k\varphi_k(x)=\sum_{k=1}^{+\infty}(c_k+\Delta c_k)\varphi_k(x)$$

is divergent at the point $x=0$. Indeed, by the hypothesis, the series $\sum\limits_{k=1}^{+\infty} c_k\varphi_k(0)$ converges to $f(0)$ while the series

$$\sum_{k=1}^{+\infty}\Delta c_k\varphi_k(0)=\delta\sqrt{\frac{2}{\pi}}\sum_{k=1}^{+\infty}\frac{\cos kx}{k}\bigg|_{x=0}=\delta\sqrt{\frac{2}{\pi}}\sum_{k=1}^{+\infty}\frac{1}{k}$$

diverges because it differs from the divergent harmonic series $\sum\limits_{k=1}^{+\infty}\frac{1}{k}$ only in the constant factor $\delta\sqrt{\frac{2}{\pi}}$ $(\delta\neq 0)$.

But nevertheless, if the exact Fourier series of a function $f(x)$ converges to $f(x)$ at a point $x\in[a,b]$ the validity of condition (6) for an arbitrarily small $\delta>0$ enables us to make the quantity

$$\left|f(x)-\sum_{k=1}^{N(\delta)}\tilde{c}_k\varphi_k(x)\right|$$

become arbitrarily small if the number of terms $N(\delta)$ is chosen in an appropriate manner. Thus, this makes it possible to obtain

at the point x, the approximate equality

$$f(x) \approx \sum_{k=1}^{N(\delta)} \tilde{c}_k \varphi_k(x) \tag{5'}$$

which is as close as desired to the exact one. Moreover, it turns out that it is unnecessary to impose any additional restrictions on the degree of smoothness of the function $f(x)$ and on the speed of convergence of its exact Fourier series at the point x. Namely, the following theorem is valid:

Theorem. *Let $\{\varphi_k(x)\}$ be an orthonormal system on an interval $[a, b]$ which satisfies the condition*

$$|\varphi_k(x)| \leqslant A = \text{const} < +\infty \quad \textit{for} \quad k = 1, 2, \ldots; \quad a \leqslant x \leqslant b \tag{13}$$

Suppose that for every δ^2 entering into the right-hand side of equality (6) the number $N(\delta)$ is chosen in such a way that the conditions

$$N(\delta) \to +\infty \quad \textit{for} \quad \delta \to 0 \tag{14$_1$}$$

and

$$\delta^2 N(\delta) \to 0 \quad \textit{for} \quad \delta \to 0 \tag{14$_2$}$$

are fulfilled. Then we have

$$\lim_{\delta \to 0} \left| f(x) - \sum_{k=1}^{N(\delta)} \tilde{c}_k \varphi_k(x) \right| = 0 \tag{15}$$

for every $x \in [a, b]$ which satisfies equality (2).

Proof. By relations (2), (4), (6), (14_1) and (14_2) and the Cauchy-Bunyakovsky inequality for sums we have

$$\begin{aligned}
\left| f(x) - \sum_{k=1}^{N(\delta)} \tilde{c}_k \varphi_k(x) \right| &= \left| \sum_{k=1}^{+\infty} c_k \varphi_k(x) - \sum_{k=1}^{N(\delta)} \tilde{c}_k \varphi_k(x) \right| \leqslant \\
&\leqslant \left| -\sum_{k=1}^{N(\delta)} \Delta c_k \varphi_k(x) \right| + \left| \sum_{k=N(\delta)+1}^{+\infty} c_k \varphi_k(x) \right| \leqslant \\
&\leqslant \sqrt{\sum_{k=1}^{N(\delta)} (\Delta c_k)^2} \sqrt{\sum_{k=1}^{N(\delta)} \varphi_k^2(x)} + \left| \sum_{k=N(\delta)+1}^{+\infty} c_k \varphi_k(x) \right| \leqslant \\
&\leqslant A \sqrt{\delta^2 N(\delta)} + \left| \sum_{k=N(\delta)+1}^{+\infty} c_k \varphi_k(x) \right|
\end{aligned} \tag{16}$$

The term $A\sqrt{\delta^2 N(\delta)}$ on the right-hand side of inequality (16) tends to zero as $\delta \to 0$ which is implied by condition (14_2), and the term $\left| \sum_{k=N(\delta)+1}^{+\infty} c_k \varphi_k(x) \right|$ tends to zero for $\delta \to 0$ because we have

condition (14_1) and the series $\sum_{k=1}^{+\infty} c_k\varphi_k(x)$ converges at the point x. The theorem has been proved.

We can now draw an important conclusion from the above proof. Let equality (2) be fulfilled at a point $x \in [a, b]$. Suppose that condition (6) imposed on the quantities Δc_k is satisfied for a given $\delta > 0$ (where Δc_k, $k = 1, 2, \ldots$, are the perturbations of the Fourier coefficients). It appears clear that to minimize the quantity $\left| f(x) - \sum_{k=1}^{N(\delta)} \tilde{c}_k\varphi_k(x) \right|$ ($\tilde{c}_k = c_k + \Delta c_k$, $k = 1, 2, \ldots$) at the point $x \in [a, b]$ we must choose the number of terms $N(\delta)$ in the partial sum $\sum_{k=1}^{N(\delta)} \tilde{c}_k\varphi_k(x)$ so that it should not be too small (because $\left| \sum_{k=N(\delta)+1}^{+\infty} c_k\varphi_k(x) \right|$ on the right-hand side of inequality (16) must be small) and too large (because the term $A\sqrt{\delta^2 N(\delta)}$ on the right-hand side of (16) must also be sufficiently small).

Every method of reconstructing a function $f(x)$, to any given accuracy, from its Fourier series with perturbed coefficients satisfying condition (6) for an arbitrarily small $\delta > 0$ is referred to as a *stable method for the summation of Fourier series with perturbed coefficients*.

Thus, the above theorem shows that if the quantity δ^2 on the right-hand side of relation (6) can be made arbitrarily small it is possible to choose $N(\delta)$ in approximate equality (5′) in an appropriate manner so as to perform stable summation of Fourier's series with perturbed coefficients.

Stable methods for the summation of Fourier's series with perturbed coefficients were developed by Tikhonov* in [16]. Tikhonov's methods make it possible to reconstruct, from a given Fourier's series with perturbed coefficients, the corresponding function $f(x)$ and also its derivatives but we shall not discuss here these more complicated questions.

---

* Tikhonov, Andrei Nikolayevich (born in 1906), a prominent Soviet mathematician.

SUPPLEMENT 1

Asymbtotic Expansions

In many problems of mathematics and mathematical physics the investigation and computation of a function $f(x)$ in the neighbourhood of a finite point x_0 or in the neighbourhood of the point at infinity* is connected with considerable difficulties. These difficulties may often be overcome by means of an *asymptotic expansion* which substitutes a simpler function for the given function $f(x)$. This simpler function is chosen in such a manner that it can be investigated and computed in an easier way than the original function $f(x)$ which it approximates to an arbitrary accuracy when x tends to x_0 or approaches infinity.

We shall begin with some examples of asymptotic expansions (§ 1) without giving general definitions, then we shall dwell in more detail on some general definitions and theorems (§ 2) and, finally, we shall illustrate Laplace's method of asymptotic representation of some integrals by deriving asymptotic formulas for the gamma function (§ 3).

§ 1. EXAMPLES OF ASYMPTOTIC EXPANSIONS

1. Asymptotic Expansions in the Neighbourhood of the Origin. Taylor's formula (e.g. see [8], Chapter 8) yields the well known asymptotic expansions for the following elementary functions:

$$
\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + o(x^n) \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^{\frac{n-1}{2}} \frac{x^{n-1}}{(n-1)!} + \\
&\qquad + o(x^n) \ \text{(where } n \text{ is odd)} \qquad (1) \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^{\frac{n}{2}} \frac{x^{n-1}}{(n-1)!} + \\
&\qquad + o(x^n) \ \text{(where } n \text{ is even)}
\end{aligned}
$$

* The expression "in the neighbourhood of the point at infinity" means "for $x \to +\infty$" or "for $x \to -\infty$" or "for $|x| \to \infty$".

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\ldots+(-1)^{n-1}\frac{x^n}{n}+o(x^n)$$

$$(1+x)^\alpha=1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\ldots \qquad (1)$$

$$\ldots+\frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}x^n+o(x^n)$$

where the symbol $o(x^n)$ designates a quantity of a higher order of smallness than x^n for $x\to 0$. Expansions (1) are used for evaluating limits of various elementary functions for $x\to 0$ when they cannot be found directly from the given analytical expressions of the functions.

2. Asymptotic Expansions in the Neighbourhood of the Point at Infinity. We now consider some asymptotic expansions in the neighbourhood of the point at infinity which are used in various applications.

In mathematical physics we are often interested in the values of the function

$$\Psi(x)=\frac{2}{\sqrt{\pi}}\int_x^{+\infty}e^{-\xi^2}\,d\xi \qquad (1)$$

for large values of the argument $x>0$.* Let us derive the asymptotic expansion of the function $\Psi(x)$ in the neighbourhood of the point at infinity, i.e. for $x\to+\infty$. Integrating by parts we find

$$\int_x^{+\infty}e^{-\xi^2}\,d\xi=\int_x^{+\infty}\frac{e^{-\xi^2}2\xi\,d\xi}{2\xi}=-\frac{e^{-\xi^2}}{2\xi}\Big|_{\xi=x}^{+\infty}-\int_x^{+\infty}\frac{e^{-\xi^2}}{2\xi^2}\,d\xi=$$

$$=\frac{e^{-x^2}}{2x}-\int_x^{+\infty}\frac{e^{-\xi^2}}{2\xi^2}\,d\xi$$

Again integrating by parts in the integral $\int_x^{+\infty}\frac{e^{-\xi^2}}{2\xi^2}\,d\xi$ etc. we obtain, after the integration by parts has been repeatedly perfor-

---

* This function can be represented in the form $\Psi(x)=\frac{2}{\sqrt{\pi}}\int_x^{+\infty}e^{-\xi^2}\,d\xi=$ $=1-\Phi(x)$ where $\Phi(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-\xi^2}\,d\xi$. The integral $\int_0^x e^{-\xi^2}\,d\xi$ is called the error function and is denoted by $\mathrm{Erf}(x)$. It plays an important role in the probability theory, theory of heat conductivity, statistical physics etc. There are various tables of its values for different values of the argument x.

med $(n-1)$ times, the following asymptotic expansion for $\Psi(x)$

$$\Psi(x)=\frac{e^{-x^2}}{\sqrt{\pi}x}\left[1+\sum_{k=1}^{n}(-1)^k\frac{1\cdot 3\ldots(2k-1)}{(2x^2)^k}+R_n(x)\right] \tag{2}$$

where

$$R_n(x)=(-1)^{n+1}\frac{1\cdot 3\ldots(2n+1)}{2^n}e^{x^2}x\int\limits_x^{+\infty}\frac{e^{-\xi^2}}{\xi^{2n+2}}d\xi \tag{3}$$

The remainder term $R_n(x)$ satisfies the following apparent inequality:

$$|R_n(x)|\leqslant\frac{1\cdot 3\ldots(2n+1)}{2^{n+1}x^{2n+2}}e^{x^2}\int\limits_x^{+\infty}e^{-\xi^2}2\xi\,d\xi=\frac{1\cdot 3\ldots(2n+1)}{2^{n+1}x^{2n+2}} \tag{4}$$

Let us designate by $O\left(\frac{1}{x^k}\right)$, for $x\to\infty$, every quantity which satisfies the relation

$$\left|O\left(\frac{1}{x^k}\right)\right|\leqslant\text{const}\cdot\frac{1}{x^k}\quad\text{for } x\to\infty$$

Then, by estimation (4), expansion (2) can be rewritten as

$$\Psi(x)=\frac{e^{-x^2}}{x\sqrt{\pi}}\left[1+\sum_{k=1}^{n}(-1)^k\frac{1\cdot 3\ldots(2k-1)}{(2x^2)^k}+O\left(\frac{1}{x^{2n+2}}\right)\right] \tag{5}$$

We have $O\left(\frac{1}{x^{2n+2}}\right)\to 0$ for $x\to+\infty$ and hence, discarding the term $O\left(\frac{1}{x^{2n+2}}\right)$, we can write the relation

$$\Psi(x)\approx\frac{e^{-x^2}}{x\sqrt{\pi}}\left[1+\sum_{k=1}^{n}(-1)^k\frac{1\cdot 3\ldots(2k-1)}{(2x^2)^k}\right]\quad\text{for } x\to+\infty \tag{6}$$

Relations (5) and (6) represent the *asymptotic expansion of the function* $\Psi(x)=\frac{2}{\sqrt{\pi}}\int\limits_x^{+\infty}e^{-\xi^2}d\xi$ *for* $x\to+\infty$, that is in the neighbourhood of the point at infinity.

Extending the summation with respect to the index k to all the values $k=1, 2, \ldots$ we obtain the asymptotic expansion

$$\Psi(x)\approx\frac{e^{-x^2}}{x\sqrt{\pi}}\left[1+\sum_{k=1}^{+\infty}(-1)^k\frac{1\cdot 3\ldots(2k-1)}{(2x^2)^k}\right] \tag{7}$$

where the series on the right-hand side of (7) is an *asymptotic serie*
This series is divergent for every value of x. Taking an arbitrar fixed and sufficiently large value of x we can easily see that whe the number k increases the modulus of the kth term of the seri first decreases and then, after its minimum value has been attaine increases and tends to infinity. But, according to inequality (4 the difference between $\Psi(x)$ and the nth partial sum of this seri satisfies the inequality

$$\left|\Psi(x)-\frac{e^{-x^2}}{x\sqrt{\pi}}\left[1+\sum_{k=1}^{n}(-1)^k\frac{1\cdot3\ldots(2k-1)}{2^kx^{2k}}\right]\right|\leqslant$$

$$\leqslant\frac{e^{-x^2}}{x\sqrt{\pi}}\cdot\frac{1\cdot3\ldots(2n+1)}{2^{n+1}x^{2n+2}} \quad (8$$

In other words, estimate (8) shows that the error arising wh the function $\Psi(x)$ is replaced by the nth partial sum of series (does not exceed, in its absolute value, the modulus of the first di carded term and thus it tends to zero fast, as $x\to+\infty$.

The repeated application of integration by parts is one of t general methods of deriving asymptotic expansions. Using th technique we can obtain asymptotic expansions for the *exponenti integral*

$$\mathrm{Ei}(x)=\int_{-\infty}^{x}\frac{e^{\xi}}{\xi}\,d\xi,\quad -\infty<x<0\quad \text{for } x\to-\infty$$

cosine integral

$$\mathrm{Ci}(x)=\int_{x}^{+\infty}\frac{\cos\xi}{\xi}\,d\xi,\quad 0<x<+\infty\quad \text{for } x\to+\infty$$

and *sine integral*

$$\mathrm{Si}(x)=\int_{0}^{x}\frac{\sin\xi}{\xi}\,d\xi,\quad -\infty<x<+\infty\quad \text{for}\quad |x|\to+\infty$$

There are various asymptotic expansions which can be obtain by means of some other elementary methods (e.g. see [3] or [6]

Let us consider a simple example. In some problems connect with studying the propagation of electromagnetic waves near t earth surface and in some other problems we encounter the functi

$$F(x)=e^{-x^2}\int_{0}^{x}e^{\xi^2}\,d\xi \quad ($$

This function can easily be expanded into a convergent power seri (see § 4, Sec. 2 of Chapter 8) but this series cannot be convenient

used for large values of x. Let us derive an asymptotic representation of $F(x)$ for $x \to +\infty$. Multiplying equality (9) by $2x$ and applying L'Hospital's rule twice for $x \to +\infty$ we obtain

$$\lim_{x \to +\infty} 2xF(x) = 1 \tag{10}$$

Consequently, we can write the following asymptotic representation for $F(x)$:

$$F(x) = \frac{1}{2x}[1 + o(1)] \quad \text{for } x \to +\infty \tag{11}$$

where the symbol $o(1)$, for $x \to +\infty$, designates a quantity which tends to zero as $x \to +\infty$. Instead of (11) we can also write

$$F(x) \approx \frac{1}{2x} \quad \text{for} \quad x \to +\infty \tag{12}$$

There are many other simple asymptotic expansions that are obtained by means of elementary methods but here we shall limit ourselves to the above examples.

§ 2. GENERAL DEFINITIONS AND THEOREMS

We shall consider functions $f(x)$, $g(x)$, ... defined on a point set M belonging to the real x-axis. For instance, as a set M we can take a finite interval, a semi-infinite interval, the whole x-axis etc.

1. **Order of Smallness. Asymptotic Equivalence.** We begin with discussing the relations of the form $f(x) = o(g(x))$ and $f(x) = O(g(x))$.

Definition 1. *If*

$$\lim_{\substack{x \to x_0 \\ x \in M}} \frac{f(x)}{g(x)} = 0 \tag{13}$$

*we say that $f(x)$ is **of higher order of smallness than** $g(x)$ on the set M for $x \to x_0$ and write*

$$f(x) = o(g(x)) \quad \textit{for} \quad x \to x_0 \quad \textit{on} \quad M \tag{14}$$

Note. Relation (13) means that for every $\varepsilon > 0$ there is $\delta = \delta(\varepsilon) > 0$ such that

$|f(x)| < \varepsilon |g(x)|$ for all $x \in M$ which satisfy the inequality

$$|x - x_0| < \delta \tag{13'}$$

Definition 2. *If there is a constant C, $0 < C < +\infty$, such that for all $x \in M$ belonging to a sufficiently small neighbourhood of x_0 the inequality*

$$\left| \frac{f(x)}{g(x)} \right| < C \tag{15}$$

*holds we say that $f(x)$ **is of the order of** $g(x)$ **on the set** M **for** $x \to x_0$ and write*

$$f(x) = O(g(x)) \quad \textit{for} \quad x \to x_0 \quad \textit{on} \quad M \tag{16}$$

Definition 2₁. *If inequality (15) is fulfilled for all $x \in M$ we say that $f(x)$ **is of the order of** $g(x)$ **on the set** M and write*

$$f(x) = O(g(x)) \quad \textit{for} \quad x \in M \tag{16'}$$

If $f(x) = o(g(x))$ for $x \to x_0$ on M, Definitions 1 and 2 indicate that we can also write $f(x) = O(g(x))$ for $x \to x_0$ but of course the relation $f(x) = O(g(x))$ does not imply, in the general case, that $f(x) = o(g(x))$.

Note. If M is an unbounded set we can similarly define the relations $f(x) = o(g(x))$ and $f(x) = O(g(x))$ for $x \to +\infty$ (or $x \to -\infty$ or $|x| \to \infty$).

In concrete problems the structure of the corresponding set M may be obvious and then we may not indicate the set M when writing the above relations.

Examples

(1) $e^x - 1 = O(x)$ for $x \to 0$,

(2) $\sin x = O(x)$ for $x \to 0$,

(3) $\cos x = O(1)$ for $x \to 0$,

(4) $\sin^2 x = o(x)$ for $x \to 0$,

(5) $x^2 = o(x)$ for $x \to 0$,

(6) $x^2 = O(x)$ for $x \to 0$,

(7) $e^{-x^2} = o(x^{-1})$ for $x \to \infty$,

(8) $F(x) = e^{-x^2} \int\limits_0^x e^{\xi^2} d\xi = O\left(\frac{1}{2x}\right)$ for $x \to +\infty$,

Now let us formulate the definition of asymptotic equivalence.

Definition 3. *We say that two functions $f(x)$ and $g(x)$ are **(asymptotically) equivalent (equal) for** $x \to x_0$ **on the set** M and write*

$$f(x) \sim g(x) \quad \textit{for} \quad x \to x_0 \quad \textit{on} \quad M \tag{17}$$

if

$$\frac{f(x)}{g(x)} \to 1 \quad \textit{for} \ x \to x_0 \ \textit{on} \ M \tag{18}$$

It is evident that instead of relation (17) we can write the relation

$$f(x) = g(x)\,[1 + o(1)] \quad \text{for} \quad x \to x_0 \quad \text{on} \quad M \tag{19}$$

which is called an asymptotic representation of $f(x)$ in the neighbourhood of the point x_0 on the set M.

The notion of functions asymptotically equivalent on an unbounded set M in the neighbourhood of the point at infinity (i.e. for $x \to +\infty$ or $x \to -\infty$ or $|x| \to \infty$) is defined similarly.

Examples

(9) $\sin x \sim x$ for $x \to 0$,

(10) $F(x) = e^{-x^2} \int\limits_0^x e^{\xi^2} d\xi \sim \frac{1}{2x}$ for $x \to +\infty$.

2. **Asymptotic Expansions of Functions.** Opening the brackets in relation (19) we get the equality

$$f(x) = g(x) + o(g(x)) \quad \text{for } x \to x_0 \quad \text{on} \quad M \tag{19'}$$

which is a simple asymptotic expansion of $f(x)$ in the neighbourhood of x_0 on the set M.

We now proceed to formulate the general definition of an asymptotic expansion which also embraces the special cases considered in § 1 of Supplement 1. We shall begin with the definitions of an *asymptotic sequence* and *asymptotic series*.

Definition 1. A finite or infinite sequence of functions $\{\varphi_n(x)\}$ defined on a set M is called an ***asymptotic sequence*** *on M for $x \to x_0$ $(x \to \infty)$ if the relations $\varphi_{n+1}(x) = o(\varphi_n(x))$ for $x \to x_0$ $(x \to \infty)$ hold for all $n = 1, 2, \ldots$.*

Here are some examples of sequences which are obviously asymptotic:

(1) $1, x, x^2, \ldots, x^n, \ldots$ for $x \to 0$,

(2) $1, x^{\lambda_1}, x^{\lambda_2}, \ldots, x^{\lambda_n} \ldots$ $(0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots)$ for $x \to 0$,

(3) $1, (x - x_0), (x - x_0)^2, \ldots, (x - x_0)^n, \ldots$ for $x \to x_0$,

(4) $1, x^{-\lambda_1}, x^{-\lambda_2}, \ldots, x^{-\lambda_n}, \ldots$ $(0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots)$ for $x \to +\infty$,

(5) $e^x, e^x x^{-\lambda_1}, e^x x^{-\lambda_2}, \ldots e^x x^{-\lambda_n}, \ldots$ $(0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots)$ for $x \to +\infty$,

(6) $1, x^{-1}, x^{-2}, \ldots, x^{-n}, \ldots$ for $x \to +\infty$.

Sequences (1), (3) and (6) are referred to as *power sequences* and sequences (2), (4) and (5) in which λ_i are real numbers that may not be integers are called *generalized power sequences*.

Definition 5. *If $\{\varphi_n(x)\}$ is an infinite asymptotic sequence (on the set M) for $x \to x_0$ $(x \to +\infty)$ the series $\sum_{n=1}^{+\infty} c_n\varphi_n(x)$ with arbitrary constant coefficients $c_1, c_2, \ldots, c_n, \ldots$ is said to be an* ***asymptotic series*** *(on the set M) for $x \to x_0$ $(x \to +\infty)$.*

Definition 6. *Let $\{\varphi_n(x)\}$ be a finite or infinite asymptotic sequence on the set M for $x \to x_0$ $(x \to \infty)$. If a function $f(x)$ defined on M satisfies the relation*

$$f(x) = \sum_{k=1}^{N} a_k\varphi_k(x) + o(\varphi_N(x)) \quad \textit{for } x \to x_0 \ (x \to \infty) \tag{20}$$

where $a_1, a_2, \ldots, a_N$ are some constants this relation is termed the ***asymptotic expansion of*** *$f(x)$ on M* ***for*** *$x \to x_0$ $(x \to \infty)$* ***up to the Nth term inclusive with respect to the sequence*** *$\{\varphi_n(x)\}$.*

Expansion (20) can also be written in the form

$$f(x) \approx \sum_{k=1}^{N} a_k\varphi_k(x) \quad \text{for } x \to x_0 \ (x \to \infty) \tag{21}$$

If asymptotic expansion (20) holds we can obviously write down the asymptotic expansions which are obtained from (20) by substituting $k = 1, 2, \ldots, N-1$ for N.

Definition 7. *Let $\sum_{k=1}^{+\infty} a_k\varphi_k(x)$ be an asymptotic series on the set M for $x \to x_0$ $(x \to \infty)$ and let asymptotic expansion* (20) *be valid for a function $f(x)$ defined on M for every $N = 1, 2, 3, \ldots$. Then this series is called an* ***asymptotic expansion of the function*** *$f(x)$* ***for*** *$x \to x_0$ $(x \to \infty)$ on M and we write*

$$f(x) \approx \sum_{k=1}^{+\infty} a_k\varphi_k(x) \quad \textit{for } x \to x_0 \quad (x \to \infty) \tag{22}$$

The asymptotic expansions considered in § 1 apparently satisfy the conditions of Definitions 6 and 7.

It should be noted that an asymptotic expansion in the neighbourhood of the point $x = 0$ $(x = x_0 \neq 0)$ can be reduced to an asymptotic expansion in the neighbourhood of the point at infinity and vice versa by a change of variable of the form $z = \frac{C}{x}$ $\left(z = \frac{C}{x - x_0}\right)$ but this may not be convenient in some practical problems.

Let us discuss the difference between an expansion of a function $f(x)$ in a functional series convergent to this function and its asymp-

totic expansion. In the former case the difference between $f(x)$ and the nth partial sum of the convergent functional series tends to zero for every fixed x as $N \to \infty$. In the latter case the difference $f(x) - \sum_{k=1}^{N} a_k \varphi_k(x)$ tends to zero, as $x \to x_0$ $(x \to \infty)$, for every fixed N, its order of smallness being higher than that of the last term in the partial sum.

The examples of asymptotic series considered in § 1 show that an asymptotic series may be convergent or divergent. Moreover, if an asymptotic series for a given function $f(x)$ is convergent this may not imply that its sum coincides with $f(x)$. For instance, we have the following simple asymptotic expansion for the function $f(x) = e^{-x}$:

$$e^{-x} \approx 0 \cdot 1 + 0 \cdot x^{-1} + \ldots + 0 \cdot x^{-n} + \ldots \qquad (\text{for } x \to +\infty) \tag{23}$$

and asymptotic series (23) is convergent but its sum is unequal to e^{-x} for all x.

For practical purposes it is important to estimate the error arising when $f(x)$ is replaced by the Nth partial sum $\sum_{k=1}^{N} a_k \varphi_k(x)$ of its asymptotic series (22), that is to estimate the remainder term $(\varphi_N(x))$ in expansion (20) for $x \to x_0$ $(x \to \infty)$.

Analytical estimation of a remainder term is often connected with considerable difficulties and therefore, in problem solving practice, we usually try to use various computational techniques based on simpler methods, which prove to be sufficient in many cases. For example, suppose we know that the absolute value of the remainder $|o(\varphi_N(x))|$ tends to zero and is a monotone decreasing function for $x \to x_0$. If we manage to compute the value of $f(x)$ at a point x lying sufficiently close to x_0 and if it turns out that the difference between this value and the value of $\sum_{k=1}^{N} a_k \varphi_k(x)$ at the same point is less than $\varepsilon > 0$ (in its modulus) we can conclude that for all values of the independent variable lying closer to x_0 than x the modulus of the difference $f(x) - \sum_{k=1}^{N} a_k \varphi_k(x)$ remains less than ε. A similar situation occurs when we have an inequality of the form $|o(\varphi_N(x))| \leqslant \psi_N(x)$ where $\psi_N(x)$ is a monotone decreasing function which tends to zero for $x \to x_0$.

If there exists an asymptotic expansion of a function $f(x)$ with respect to a given asymptotic sequence $\{\varphi_n(x)\}$ this expansion is uniquely specified by $f(x)$. Namely, we shall prove the following theorem.

Theorem 1. Let every member of an asymptotic sequence $\{\varphi_n(x)\}$ *be different from zero for all* x *belonging to a sufficiently small neighbourhood of* x_0 *(or for* $x \to \infty$*) and let an asymptotic expansion of form* (20) *hold for a function* $f(x)$. *Then the coefficients* a_k, $k = 1, 2, \ldots, N$, *of this expansion are uniquely determined by the formulas*

$$a_n = \lim_{\substack{x \to x_0 \\ (x \to \infty)}} \frac{f(x) - \sum\limits_{k=1}^{n-1} a_k \varphi_k(x)}{\varphi_n(x)} \quad \text{for } n = 1, 2, \ldots, N \tag{24}$$

Proof. Replacing N by $n \leqslant N$ in relation (20) we put it in the form

$$f(x) = \sum_{k=1}^{n-1} a_k \varphi_k(x) + a_n \varphi_n(x) + o(\varphi_n(x)), \quad 1 \leqslant n \leqslant N$$

and find

$$a_n = \frac{f(x) - \sum\limits_{k=1}^{n-1} a_k \varphi_k(x)}{\varphi_n(x)} + \frac{o(\varphi_n(x))}{\varphi_n(x)}, \quad 1 \leqslant n \leqslant N$$

which implies the validity of formulas (24). The theorem has been proved.

It should be noted that the converse of the above theorem is not true. Namely, a function $f(x)$ is not uniquely determined by its asymptotic expansions, that is there may exist different functions having the same asymptotic expansion. For example, the functions $f(x) = e^{-x}$ and $g(x) = 0$ have the same asymptotic expansion (23) with respect to the sequence $1, x^{-1}, x^{-2}, \ldots, x^{-n}, \ldots$ for $x \to +\infty$.

Definition 8. Two functions $f(x)$ *and* $g(x)$ *are said to be* ***asymptotically equivalent (equal)*** *with respect to a given asymptotic sequence* $\{\varphi_n(x)\}$ *for* $x \to x_0$ $(x \to \infty)$ *if the relation*

$$f(x) - g(x) = o(\varphi_n(x)) \quad \text{for } x \to x_0 \ (x \to \infty) \tag{25}$$

is fulfilled for all n.

It appears obvious that two functions $f(x)$ and $g(x)$ possessing the same asymptotic expansion with respect to an asymptotic sequence are equivalent (relative to this sequence).

We can easily show that for two functions $f(x)$ and $g(x)$ to have the same coefficients of their asymptotic expansions with respect to one and the same asymptotic sequence $\{\varphi_n(x)\}$ for $x \to x_0$ $(x \to \infty)$ it is necessary and sufficient that these functions be asymptotically equal with respect to the sequence $\{\varphi_n(x)\}$ for $x \to x_0$ $(x \to \infty)$.

We now discuss some operations on asymptotic expansions.

If we have asymptotic expansions

$$f(x) \approx \sum_{k=1}^{+\infty} a_k \varphi_k(x) \quad \text{and} \quad g(x) \approx \sum_{k=1}^{+\infty} b_k \varphi_k(x) \quad \text{for} \quad x \to x_0 \ (x \to \infty) \tag{26}$$

we can obviously write down the asymptotic expansion

$$\alpha f(x) + \beta g(x) \approx \sum_{k=1}^{+\infty} (\alpha a_k + \beta b_k) \varphi_k(x) \quad \text{for} \quad x \to x_0 \ (x \to \infty) \tag{27}$$

which is valid for any constants α and β.

In the general case it is not allowable to multiply the asymptotic expansions of two functions $f(x)$ and $g(x)$ with respect to one and the same asymptotic sequence $\{\varphi_n(x)\}$ because it may turn out that the products $\varphi_m(x)\,\varphi_n(x)$ cannot be arranged into an asymptotic sequence.

Let us dwell on the question of term-by-term integration of asymptotic expansions.

Theorem 2. Let a sequence $\{\varphi_n(x)\}$ of positive functions of a real variable x defined on an interval $a < x < b$ be an asymptotic sequence for $x \to b - 0$. Suppose we have an asymptotic expansion

$$f(x) \approx \sum_{k=1}^{+\infty} C_k \varphi_k(x) \quad \textit{for} \ \ x \to b - 0^* \tag{28}$$

If the integrals

$$\int_x^b f(\xi)\,d\xi \quad \textit{and} \quad \int_x^b \varphi_k(\xi)\,d\xi, \quad k = 1, 2, \ldots \tag{29}$$

are convergent the asymptotic expansion

$$\int_x^b f(\xi)\,d\xi \approx \sum_{k=1}^{+\infty} C_k \int_x^b \varphi_k(\xi)\,d\xi \quad \textit{for} \quad x \to b - 0 \tag{30}$$

is also valid.

Proof. The relation $\varphi_{n+1}(x) = o(\varphi_n(x))$ means that for every $\varepsilon > 0$ we have the inequality $|\varphi_{n+1}(x)| < \varepsilon\, |\varphi_n(x)|$ provided that n is sufficiently large.** The functions $\varphi_k(x)$, $k = 1, 2, \ldots$, being positive, we can drop the sign of modulus and write

$$\varphi_{n+1}(x) < \varepsilon \varphi_n(x) \quad \text{for} \ \ x \to b - 0 \tag{$*$}$$

* Here b is a finite number or $+\infty$.

** See the note after Definition 1.

for any $\varepsilon > 0$. Integrating (*) from x to b (which is permissibl since the integrals are convergent) we see that

$$0 < \int_x^b \varphi_{n+1}(\xi)\, d\xi < \varepsilon \int_x^b \varphi_n(\xi)\, d\xi \tag{31}$$

for any $\varepsilon > 0$ and $x \to b-0$ whence it follows that $\left\{\int_x^b \varphi_n(\xi)\, d\xi\right\}$ is an asymptotic sequence for $x \to b-0$. To prove relation (30 we must show that

$$\int_x^b f(\xi)\, d\xi \approx \sum_{k=1}^{N} C_k \int_x^b \varphi_k(\xi)\, d\xi \quad \text{for} \quad x \to b-0 \tag{30'}$$

for every $N = 1, 2, 3, \ldots$. By the positivity of the function: $\varphi_n(x)$, $n = 1, 2, \ldots$, relation (20) (whose validity is implied by (28)) can be rewritten in the form

$$\left| f(x) - \sum_{k=1}^{N} a_k \varphi_k(x) \right| < \varepsilon \varphi_N(x) \quad \text{for} \quad x \to b-0$$

for any $\varepsilon > 0$ where $N = 1, 2, 3, \ldots$. Consequently, we have

$$\left| \int_x^b f(\xi)\, d\xi - \sum_{k=1}^{N} a_k \int_x^b \varphi_k(\xi)\, d\xi \right| \leqslant$$

$$\leqslant \int_x^b \left| f(\xi) - \sum_{k=1}^{N} a_k \varphi_k(\xi) \right| d\xi \leqslant \varepsilon \int_x^b \varphi_N(\xi)\, d\xi \quad \text{for} \quad x \to b-0$$

and any $\varepsilon > 0$ for every $N = 1, 2, 3, \ldots$. But this exactly means that asymptotic expansion (30) holds for every $N = 1, 2, 3, \ldots$, which is what we set out to prove.

The assertion below is a direct consequence of Theorem 2:

If we have a power asymptotic expansion

$$f(x) \approx \sum_{k=0}^{+\infty} a_k x^{-k} \quad \textit{for} \quad x \to +\infty \tag{32}$$

and the integral $\int_x^{+\infty} [f(\xi) - a_0 - a_1\xi^{-1}]\, d\xi$ *converges, the asymptotic expansion*

$$\int_x^{+\infty} [f(\xi) - a_0 - a_1\xi^{-1}]\, d\xi \approx \sum_{k=2}^{+\infty} \frac{a_k}{-k+1} x^{-k+1} \tag{33}$$

is also valid.

In the general case an asymptotic expansion cannot be differentiated termwise. But there are some special cases of asymptotic expansions when term-by-term differentiation is permissible. For instance, let $f(x)$ have a power asymptotic expansion of form (32). Suppose that its derivative $f'(x)$ also possesses a power asymptotic expansion in which there are no terms containing x^0 and x^{-1}:

$$f'(x) \approx -\frac{a_1^*}{x^2} - \frac{2a_2^*}{x^3} - \ldots - \frac{na_n^*}{x^{n+1}} - \ldots \quad \text{for } x \to +\infty \tag{34}$$

Integrating (34) we obtain, by the foregoing assertion, the relation

$$f(x) \approx f(+\infty) + \frac{a_1^*}{x} + \frac{a_2^*}{x^2} + \ldots + \frac{a_n^*}{x^n} + \ldots \quad \text{for } x \to +\infty \tag{35}$$

But since the function $f(x)$ uniquely specifies its asymptotic series with respect to the sequence $1, x^{-1}, x^{-2}, x^{-3}, \ldots, x^{-k}, \ldots$ expansion (35) must coincide with expansion (32) and thus we have the equalities $f(+\infty) = a_0$, $a_1^* = a_1, \ldots, a_k^* = a_k, \ldots$. Substituting the above values into (34) we obtain

$$f'(x) \approx -\frac{a_1}{x^2} - \frac{2a_2}{x^3} - \ldots - \frac{na_n}{x^{n+1}} - \ldots \quad \text{for } x \to +\infty \tag{34'}$$

where the latter asymptotic expansion can be directly obtained from asymptotic relation (32) by termwise differentiation.

We shall limit ourselves to this short review of some general properties of asymptotic expansions. In § 3 we shall describe an important method of constructing asymptotic expansions of some integrals.

§ 3. LAPLACE METHOD FOR DERIVING ASYMPTOTIC EXPANSIONS OF SOME INTEGRALS

Let it be required to derive an asymptotic representation of an integral

$$J(t) = \int_a^b f(x, t)\,dx \quad \text{for } t \to +\infty \tag{36}$$

under the assumption that, for large values of t, the integrand has a sharp extremum in a neighbourhood of a point $x = x_0$ and that its modulus is very small outside this neighbourhood. Then it may turn out that the integral taken over this neighbourhood of the point x_0 will be almost equal to the integral taken from a to b for large values of t. If, in addition, it is possible to replace the function $f(x, t)$ in this neighbourhood, to a sufficient degree of accuracy, by a simple function which can easily be integrated, and if the difference between the original integral (36) and the integral of this auxiliary function tends to zero as $t \to +\infty$, we can thus derive

an asymptotic expansion of integral (36) for large values of t. This is the general idea of the method suggested by Laplace for deriving asymptotic expansions of integrals of the above type.

As an illustrative instance, we shall apply this method to deriving an asymptotic expansion of the gamma function

$$\Gamma(t+1)=\int_{0}^{+\infty} e^{-u}u^{t}\,du \tag{37}$$

for $t\to+\infty$. The asymptotic representation which we shall thus obtain for $\Gamma(t+1)$ for large values of t yields, in the case when $t=n>0$ is an integer, the so-called Stirling* formula which gives an asymptotic representation of $n!$ for large values of n.

Performing the substitution $u=t(1+x)$ we reduce integral (37) to the form

$$\Gamma(t+1)=e^{-t}t^{t+1}\int_{-1}^{+\infty}[e^{-x}(1+x)]^{t}\,dx=e^{-t}t^{t+1}\int_{-1}^{+\infty}e^{th(x)}\,dx \tag{38}$$

where $h(x)=-x+\ln(1+x)$. The function $e^{-x}(1+x)=e^{-x+\ln(1+x)}=e^{h(x)}$ attains its maximum at the point $x=0$ at which $h(x)$ has its only maximum. Indeed, the derivative $h'(x)=-1+\frac{1}{1+x}$ is positive for $-1<x<0$ and negative for $0<x<+\infty$. At the point of maximum of $h(x)$ we have $h(0)=0$ and hence the function $h(x)$ is negative for all the values of x belonging to the intervals $-1<x<0$ and $0<x<+\infty$. Therefore, the function $e^{th(x)}$ tends to zero as $t\to+\infty$ for $-1<x<0$ and $0<x<+\infty$. Consequently, it is advisable to try to apply Laplace's method.

Taking a sufficiently small $\delta>0$ $(\delta<1)$ we represent the integral in question as the sum

$$\int_{-1}^{+\infty}e^{th(x)}\,dx=\int_{-1}^{-\delta}e^{th(x)}\,dx+\int_{-\delta}^{\delta}e^{th(x)}\,dx+\int_{\delta}^{+\infty}e^{th(x)}\,dx \tag{39}$$

To estimate each integral on the right hand side of equality (39) we investigate $h(x)$ on the intervals $-1<x\leqslant-\delta$, $-\delta\leqslant x\leqslant\delta$ and $\delta\leqslant x<+\infty$. Expanding $\ln(1+x)$ into Taylor's series in the interval $-\delta\leqslant x\leqslant\delta$ we obtain

$$h(x)=-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\frac{x^5}{5}-\ldots=$$
$$=\frac{x^2}{2}\left[-1+\frac{2x}{3}-\frac{2x^2}{4}+\frac{2x^3}{5}-\ldots\right]$$

---

* Stirling, James (1692-1770), an English mathematician.

The series in the square brackets satisfies the conditions of Leibniz' test because $|x| < \delta < 1$. Therefore if we discard all the terms beginning with $-\frac{2x^2}{4}$ this results in an error whose modulus does not exceed $\frac{2x^2}{4}$. For sufficiently small $\delta > 0$ we have $\frac{2x^2}{4} < \frac{2|x|}{3}$ for all x belonging to the interval $-\delta \leqslant x \leqslant \delta$. Consequently, we can write

$$\frac{x^2}{2}\left[-1-\frac{4}{3}\delta\right] \leqslant h(x) \leqslant \frac{x^2}{2}\left[-1+\frac{4}{3}\delta\right] \quad \text{for } -\delta \leqslant x \leqslant \delta,\ \frac{4}{3}\delta < 1 \tag{40}$$

The function $h(x)$ increases on the interval $-1 \leqslant x \leqslant -\delta$ and assumes its greatest value $h(-\delta) < 0$ at the point $x = -\delta$. On the interval $\delta \leqslant x < +\infty$ the function $h(x)$ decreases and its greatest value $h(\delta) < 0$ is attained at the point $x = \delta$. Thus,

$$h(x) \leqslant h(-\delta) < 0 \quad \text{for} \quad -1 < x \leqslant \delta \tag{41}$$

and

$$h(x) \leqslant h(\delta) < 0 \quad \text{for} \quad \delta \leqslant x < +\infty \tag{42}$$

Now let us estimate the integrals on the right-hand side of (39). By inequality (41), we have

$$\int_{-1}^{-\delta} e^{th(x)}\,dx \leqslant \int_{-1}^{-\delta} e^{th(-\delta)}\,dx = e^{th(-\delta)}(1-\delta) = O(e^{th(-\delta)}) \to 0 \tag{43}$$

for $t \to +\infty$. Furthermore, for $\delta \leqslant x < +\infty$ and $t > 1$ we have, by (42), the inequalities $th(x) \leqslant th(\delta)$ and $th(x) \leqslant h(x)$. Adding them up we get the relation

$$th(x) \leqslant \frac{1}{2}[th(\delta) + h(x)] \quad \text{for } t > 0 \quad \text{and} \quad \delta \leqslant x < +\infty \tag{44}$$

Therefore

$$\int_{\delta}^{+\infty} e^{th(x)}\,dx \leqslant \int_{\delta}^{+\infty} e^{\frac{1}{2}(th(\delta)+h(x))}\,dx =$$

$$= e^{\frac{1}{2}h(\delta)} \int_{\delta}^{+\infty} e^{\frac{1}{2}h(x)}\,dx = O(e^{\frac{1}{2}th(\delta)}) \to 0 \tag{45}$$

for $t \to +\infty$. Finally, by inequality (40), we can write

$$\int_{-\delta}^{\delta} e^{-t\frac{x^2}{2}\left[1+\frac{4}{3}\delta\right]}\,dx < \int_{-\delta}^{\delta} e^{th(x)}\,dx < \int_{-\delta}^{\delta} e^{-t\frac{x^2}{2}\left[1-\frac{4}{3}\delta\right]}\,dx \tag{46}$$

On the basis of estimates (44) and (45) we obtain the relation

$$\int_{-\delta}^{\delta} e^{-t\frac{x^2}{2}\left[1\pm\frac{4}{3}\delta\right]} dx = \int_{-\infty}^{+\infty} e^{-t\frac{x^2}{2}\left[1\pm\frac{4}{3}\delta\right]} dx + O\left(e^{-\alpha(\delta)t}\right) \qquad (47$$

for $t \to +\infty$ where $\alpha(\delta)$ is a positive constant dependent on and independent of t. Now let us evaluate the integral

$$\int_{-\infty}^{+\infty} e^{-t\frac{x^2}{2}\left[1\pm\frac{4}{3}\delta\right]} dx$$

Performing the change of variable $\xi = x\left[\frac{t}{2}\left(1 \pm \frac{4}{3}\delta\right)\right]^{\frac{1}{2}}$ we obtai

$$\int_{-\infty}^{+\infty} e^{-t\frac{x^2}{2}\left[1\pm\frac{4}{3}\delta\right]} dx = (2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} \left(1 \pm \frac{4}{3}\delta\right)^{-\frac{1}{2}} \qquad (48$$

From (47) and (48) it follows, for $t \to +\infty$, that

$$\int_{-\delta}^{\delta} e^{-t\frac{x^2}{2}\left[1\pm\frac{4}{3}\delta\right]} dx = (2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} \left(1 \pm \frac{4}{3}\delta\right)^{-\frac{1}{2}} + O\left(e^{-\alpha(\delta)t}\right) \qquad (49)$$

It is obvious that

$$\left(1 + \frac{4}{3}\delta\right)^{-\frac{1}{2}} = 1 - \varepsilon_1(\delta) \quad \text{and} \quad \left(1 - \frac{4}{3}\delta\right)^{-\frac{1}{2}} = 1 + \varepsilon_2(\delta)$$

where $\varepsilon_1(\delta)$ and $\varepsilon_2(\delta)$ are positive and tend to zero for $\delta \to 0$. If $\delta > 0$ is fixed we have, for sufficiently large $t > 0$, the inequality

$$|O\left(e^{-\alpha(\delta)t}\right) < \min\left(\varepsilon_1(\delta),\ \varepsilon_2(\delta)\right)(2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}}$$

since the quantity $-\alpha(\delta)$ is negative. Therefore from (46) and (49) we deduce the relation

$$(2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} |1 - 2\varepsilon_1(\delta)| < \int_{-\delta}^{\delta} e^{th(x)} dx < (2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} |1 + 2\varepsilon_1(\delta)| \qquad (50)$$

for sufficiently large $t > 0$. Taking into account relations (43) and (45) and the fact that the inequalities

$$|O\left(e^{th(-\delta)}\right)| < (2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} \min\left(\varepsilon_1(\delta),\ \varepsilon_2(\delta)\right)$$

and

$$|O\left(e^{th(\delta)}\right)| < (2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} \min\left(\varepsilon_1(\delta),\ \varepsilon_2(\delta)\right)$$

hold for sufficiently large $t > 0$ we arrive at the inequality

$$(2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} (1 - 3\varepsilon_1(\delta)) < \int_{-1}^{+\infty} e^{th(x)}\, dx < (2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} (1 + 3\varepsilon_2(\delta)) \quad (51)$$

which holds for all sufficiently large $t > 0$. Consequently, since $\varepsilon_1(\delta)$ and $\varepsilon_2(\delta)$ are arbitrarily small if δ is sufficiently small we can write, taking into account the definition of the symbol $o(1)$, the relation

$$\int_{-1}^{+\infty} e^{th(x)}\, dx = (2\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} [1 + o(1)] \quad \text{for } t \to +\infty \quad (52)$$

Substituting (52) into (38) we finally obtain the sought-for asymptotic representation of $\Gamma(t+1)$ for $t \to +\infty$:

$$\Gamma(t+1) = e^{-t} t^{t+\frac{1}{2}} (2\pi)^{\frac{1}{2}} [1 + o(1)], \quad t \to +\infty \quad (53)$$

Putting $t = n$ where n is a positive integer we derive from relation (53) **Stirling's formula**

$$n! = e^{-n} n^{n+\frac{1}{2}} (2\pi)^{\frac{1}{2}} [1 + o(1)] \quad \text{for } n \to +\infty \quad (54)$$

which is widely used in mathematics and its applications.

The Laplace method makes it possible to obtain asymptotic expansions of a more general type. It can also be applied to multiple integrals (see [3]). Laplace's method is also generalized to the case of integrals of functions of a complex variable. This generalization is known as the **saddle-point method** which, like Laplace's method, is used in various divisions of mathematics and mathematical physics. On the saddle-point method we refer the reader to [3] and [12].

SUPPLEMENT 2

On Universal Digital Computers

This supplement provides an introduction to modern digital computers, their operation and use. It cannot be regarded as a systematic account of the computer theory and programming methods, and for greater detail we refer the reader to special books.

§ 1. COMPUTERS

1. Introduction. Many problems of modern science and engineering require extensive calculations to obtain results of practical importance. The amount of work may be so large that the calculations either cannot be carried out manually or take so much time that the result becomes useless. For example, it makes no sense to forecast the next day's weather by applying a method that takes a month of computational work.

The number of problems requiring large-scale calculations as well as prompt answers has recently increased in the light of such technical needs as automation of manufacturing processes.

Some technical devices which facilitate computations were invented long ago but the last two decades are connected with a radical turn in this field owing to the appearance of *high speed computing machines* based on electronics.

The new technique has achieved a marvellous success in a short time. Modern computers can do hundreds of thousands of arithmetical operations per second. This makes it possible to succeed in solving such problems that could not even be stated before.

Computers not only increased the range of mathematical applications but also influenced the development of mathematics itself. In mathematical logic and numerical analysis there have arisen new problems and new trends. The computer theory and programming (see § 3) are very important fields of modern mathematics.

Nowadays computers are used almost everywhere. Therefore specialists in various branches of science, physicists in particular, should be familiar with the mode of operation of these machines, with their properties and capabilities.

2. Basic Types of Computer. Computing machines are divided into two basic classes: machines of discrete operation referred to

as **digital** and machines of continuous operation called **analogue**. A digital computer operates with numbers represented in a positional number system. An analogue computer represents variables by means of some physical processes and quantities (such as electric currents, voltages, mechanical displacements and so on) that may vary continuously. In analogue computers only the final results take the digital form. Analogue computers are widely applied (mostly when high accuracy is not required) but in modern computational mathematics they are of less importance than digital machines.

In what follows we shall deal with digital computers to which the general term "computer" will be applied.

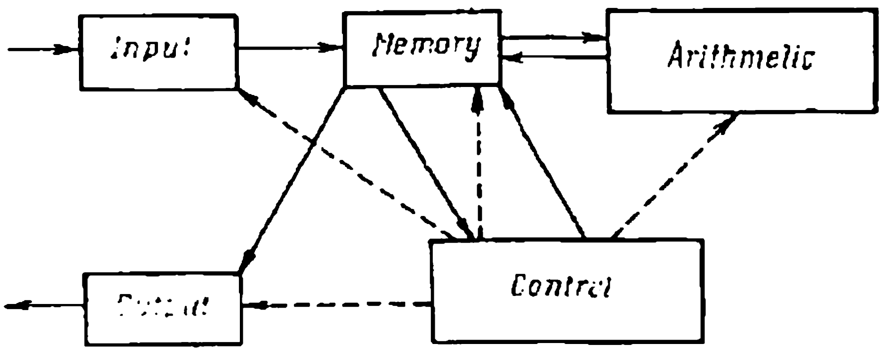
There are two types of digital computers: **special purpose** and **general purpose**. Special purpose computers are designed to solve a specific class of problems. A computer is said to be general purpose or universal when it can be used to solve a wide variety of problems, arithmetical as well as logical (for instance, the translation of languages). It is the universal computers that we are going to consider here.

Versatility is a great advantage of universal computers. However, it should be noted that the computer itself can only perform a restricted number of elementary operations (see § 2), and therefore to solve a problem on a computer we must reduce it to a sequence of elementary steps, that is a routine corresponding to the problem must be prepared for the machine. The process of preparing a routine (programming) may be rather complicated. Some general notions related to programming and examples of elementary routines are given in § 3 of this supplement.

3. Principal Components of a Computer and Their Functions. As has been mentioned, every computer can perform some elementary operations, arithmetical and logical. Furthermore, facilities must be provided for performing some additional operations, namely for entering **initial data** and **instructions** into the machine, storing this information and intermediate results and withdrawing final results from the computer. Accordingly, every universal digital computer includes, irrespective of its construction, the following functional units:

1. *Input unit.*
2. *Storage unit (memory).*
3. *Arithmetic unit.*
4. *Control unit.*
5. *Output unit.*

A block diagram of a typical computing machine system is shown below.



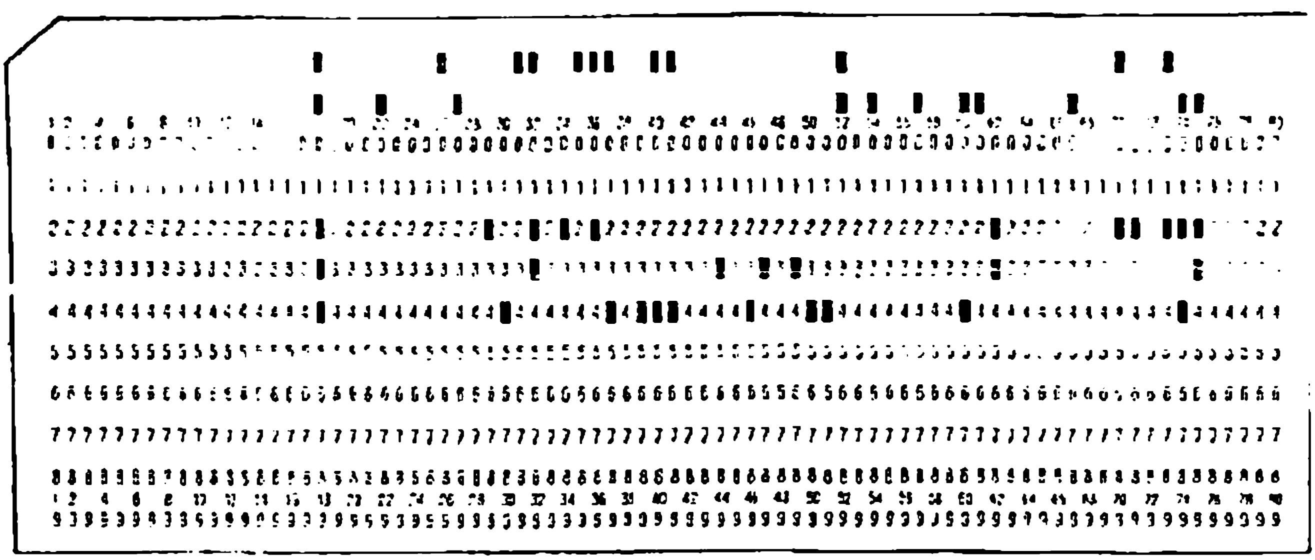
Direction of information flow is shown by continuous arrows and that of control signals by dotted ones.

Let us briefly consider the function of each unit.

1. **Input** is a device used for transferring initial data and instructions into a computer. Information must be prepared in a form intelligible to the machine. In general, it processes some physical medium such as magnetic tape, punch tape and punch cards.

On a magnetic tape data are stored as small magnetized spots arranged in column form across the width of the tape.

On a punch tape and punch cards storage is in the form of holes punched at definite places. A punch card is shown below. Such cards are widely used.



2. **Storage** (**memory**) is the section in which instructions, initial data and intermediate results are stored and from which they can be taken. A storage is a set of **locations**, **cells**, each of which is identified by a number. Each location is of a fixed digit capacity and can be used for storing numbers as well as instructions. As is seen from the block diagram, the storage unit is directly connected with the other units. From the input device it accepts the initial data and the set of instructions (the **program**) corresponding to the problem which is to be solved. The storage sends numbers to the arith-

metic unit where arithmetical operations are carried out. The results come back to the storage.

Complicated problems require storage devices of great capacity. At the same time it is important that the stored data could be rapidly read out. To satisfy both conditions storage devices are usually made of two blocks: **internal** and **external memory**. An external (auxiliary) storage holds much larger amount of information than the main, internal, storage but its access time is greater. The external storage is not connected directly to the arithmetic unit; if necessary it feeds information to the internal storage and only then this information can be processed by the arithmetic unit.

3. **Arithmetic unit** is a part of a computer in which arithmetic operations are performed. The set of these operations will be specified below.

4. **Control** is the section which interprets the instructions involved in a routine (i.e. converts them into pulses) and then sends appropriate signals to other machine blocks. The control unit determines, in accordance with a given routine, the operation of all other parts of the computer.

Since no human operator can achieve the speed at which a computer works a control unit must be automatic; this is one of the main principles of computer organization.

5. **Output unit** is meant for accepting, in a suitable form, the solution of a problem and some intermediate results that are of interest. A few output devices are card punches, paper tape punches, printers and magnetic tape units.

4. Number Systems Used in Computers. When we operate on numbers (no matter whether we utilize a computer or not) we have to represent them in a certain system of notation, i.e. in a number system. The system practised on the greatest scale nowadays is the **decimal system** in which any numerical quantity is represented as a sequence of coefficients in the successive powers of ten. For example, the decimal expression 2548 denotes the number

$$2\cdot 10^3+5\cdot 10^2+4\cdot 10^1+8\cdot 10^0$$

The decimal number system uses the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9; operations on numbers comply with the well known rules.

Any other positive integer except unity may also be used as a base of a number system.* Logically, the simplest is the **binary system** in which every number is represented as a combination of powers of two. For example, $13=1\cdot 2^3+1\cdot 2^2+0\cdot 2^1+1\cdot 2^0$

* Indeed, there were nations that used non-decimal number systems. From the mathematical point of view the decimal system has no special advantages, and its general use stems from the fact that man has ten fingers.

and hence the decimal number "thirteen" is represented as the binary number 1101. This notation uses only two digits 0 and 1, the number two being the unity of the next (to the left) position.

The reader might well ask at this point which system of notation is the best suited for electronic computers. Note that if we use a number system with base p the corresponding digits can have one of the values $0, 1, 2, \ldots, p-1$ (for instance, the decimal system uses 10 digits, the binary system uses 2 digits etc.). In order to be able to fix p different digits a machine must involve some devices that have p stable states, each representing finite digit. The speed with which a modern computer works (as a rule, hundreds of thousands of operations a second) makes it impossible to use any mechanical device for fixing numbers. On the other hand such speed can be easily achieved by electronic circuits which are practically inertialess. Most electronic elements (valves, transistors etc.) are bi-stable. For instance, a valve assumes either of two stable states: "on" (the current flows through it) and "off" (the current does not flow through it). Owing to these features of electronic devices, the binary number system proves to be the most applicable for modern calculating equipment.

The binary system has also the advantage of simpler arithmetic. For example, the "multiplication table", in the binary notation, consists of the following four items:

$$0\cdot 0 = 0, \qquad 0\cdot 1 = 0,$$

$$1\cdot 0 = 0, \qquad 1\cdot 1 = 1$$

The binary system has some disadvantage because it requires converting initial data written to the base ten to the equivalent numbers written to the base two and converting computed results back to the decimal form. This operation is however not complicated and may easily be automatized.

Besides the binary system, computers also use the octal system (based on a radix of eight) and the binary coded decimal notation (binary-decimal system). In the latter a number is first written in the decimal notation and then each decimal digit is represented by the corresponding binary number. For instance, the number

$$5386$$

is represented in the binary-decimal notation as

$$0101,\ 0011,\ 1000,\ 0110$$

It is clear that each decimal digit (i.e. the digits $0, 1, 2, \ldots, 9$) can be represented by a four-digit binary number.

There are computers with elements having three stable states (for example, the current flows through an element, the current flows

in the opposite direction and the current does not flow). Arithmetic in these machines is based on the **ternary number system**.

5. Representing Numbers Within a Computer. Any computer operates with quantities having a definite (for each machine) number of digits. If a number is shorter than the location in a given machine, zeros are put to the left of significant digits. If a number is longer it must be rounded off by deleting less significant digits. The number of binary places in each location limits the precision of representing the results and thus restricts the accuracy of computations.

Since we deal with negative quantities as well as with positive ones, the computer must contain some means of distinguishing between them. A certain binary position is usually assigned for this purpose and the codes 0 and 1 are interpreted, respectively, as positive and negative signs of the quantities.

Moreover, since we have to deal with mixed numbers the machine must be able to separate integral and fractional parts by a "binary point". The position of the radix point may either vary in the course of calculation (**floating-point machines**) or be constant (**fixed-point machines**). In the latter case the integral part of any numerical quantity is expressed by a predetermined number of digits. All quantities occurring in a problem to be solved on a fixed-point computer must be converted into the desired range of magnitude by means of "**scale-factors**" specific for each problem. This makes fixed-point machines less convenient than floating-point ones; however, their logic and hardware are simpler.

§ 2. BASIC OPERATIONS EXECUTED BY A COMPUTER. INSTRUCTIONS

1. Types of Operation. We have already mentioned that any computer is designed to perform a certain restricted number of basic operations. Although this number can be still reduced, it would be inconvenient for the programmer and user. On the other hand, the larger the number of different elementary operations, the more complicated the construction of the machine. A compromise must therefore be reached between the claims of the designer and of the programmer; the point at which the balance is struck varies from one machine to another, every computer still being able to carry out the following operations:

(1) *arithmetical operations*;

(2) *additional computational operations*;

(3) *logical operations*;

(4) *transfer operations* (including *conditional transfer of control*);

(5) *input and output operations.*

One point must be emphasized at the outset. We assume all operands to be stored in the memory locations, each location being identified with an **address**, i.e. a certain serial number. When

an operation on two quantities is performed the following must be indicated:

(1) the addresses of the operands;

(2) the command to be executed (addition, multiplication etc.);

(3) the address for storing the result.

Consequently, each *instruction* that causes a computer to perform an operation contains three addresses of the locations involved in the operation and specifies the operation which is to be performed. In other words, such an instruction is of the form*

Operation	1st address	2nd address	3rd address

It is important to note here that it is not the operands but their addresses that are indicated in the instruction. This enables us to prepare a program (a sequence of instructions) without knowing the specific values of the quantities to be dealt with.

Now let us describe the basic operations.

2. **Arithmetical Operations.** Arithmetical operations are of the following four types:

(1) *Addition*. This operation reads: "Take the number stored in location α, add it to the number stored in location β and store the result in location γ." Symbolically:

Addition	α	β	γ

(2) *Multiplication*: "Multiply the number in location α by the number in location β and store the result in location γ." Symbolically:

Multiplication	α	β	γ

(3) *Subtraction*: "Subtract the number in location β from the number in location α and store the result in location γ." Symbolically:

Subtraction	α	β	γ

* For the sake of simplicity, we consider computers referred to as three-address machines. There exist machines that involve one-, two-, or four-address instructions but we shall not discuss them.

(4) *Division*: "Divide the number in location α by the number in location β and store the result in location γ." Symbolically:

Division	α	β	γ

3. Additional Computational Operations. The set of these operations may vary from one machine to another. Here are some examples:

(1) *Maximum*: "From two numbers stored in locations α and β take the greater one and place it in location γ", that is

Maximum	α	β	γ

(2) *Minimum*: is defined similarly.

(3) *Magnitude*: "Take the absolute value of the number in location α and store it in location γ", i.e.

Absolute value	α		γ

The latter operation (and some other) involves only two (and not three) addresses.

The greater the number of various additional operations, the easier the process of programming. Many modern computers have built-in instructions for square root operation, sine evaluation and so on, although these operations are in fact reduced to certain combinations of a limited variety of elementary computing steps (see § 3).

4. Logical Operations. These operations on numbers are performed on a digit-by-digit basis without carry. A few examples will illustrate this type of operation.

(1) *Logical addition*. It is an operation in which the numbers placed in locations α and β are added *bit-to-bit**, i.e. each digit in α is added to the corresponding digit in β according to the following rules:

$$0 + 0 = 0, \quad 0 + 1 = 1, \quad 1 + 0 = 1, \quad 1 + 1 = 0$$

The result is placed in location γ. We denote this operation symbolically as

Bit-to-bit addition	α	β	γ

* The term *bit* means binary unit (of information) or binary digit.

(2) *Comparison.* This operation is concerned with the determination of similarity or dissimilarity of the corresponding digits in the numbers stored in locations α and β. If the digits in a certain place in α and β are alike the result in this place at γ is equal to unity. if otherwise. the result is equal to zero.* For example, the result of comparing the numbers

1 0 1 1 0 1 1 0 1
1 1 0 1 0 0 1 0 0

is the number

1 0 0 1 1 0 1 1 0

Let us symbolically denote this operation by

Comparison	α	β	γ

(3) *Logical Negation.* The operation is performed as follows: if the number stored in location α contains zero in a certain position the corresponding binary place in location γ contains unity and vice versa. Symbolically:

Logical negation	α		γ

The set of available logical operations depends upon the type of a computer.

5. **Input and Output Operations.** These are the following operations: input, **writing** (transferring a number from the internal storage to the external storage), **reading** (transferring a number in the reverse direction), **printing** and **stopping**.

(1) *Input instruction* is denoted as

Input	$n-1$	$\alpha+1$	

It means: "Transfer n numbers (or instructions) from the input device (e.g. punch cards or punch tape) into n memory locations $\alpha+1$, $\alpha+2$, . . ., $\alpha+n$."

(2) *Print.* This is an instruction of the form

Print	α	$n-1$	

* Some computers use the opposite rule, i.e. 0 denotes the fact of coincidence and 1 of discrepancy.

which means "Print (in the decimal system) the numbers stored in n consecutive locations beginning with location α."

(3) *Stop*. This is an instruction of the form

Stop			

which causes the computer to stop working.

External storage devices are used in large-scale computations requiring long programs and large amounts of initial data. We shall not dwell on them here.

6. **Transfer of Control.** We have pointed out that it is not the operands but their addresses that are indicated in an instruction. This enables us to plan the whole procedure for solving a problem before starting computations. However, in a computational process there may occur a situation in which the further course of computation depends upon the result we have obtained at a certain stage. For example, if a quadratic equation is being solved the course of computing depends on the sign of its discriminant. If some alternative courses of action must be taken depending on a relationship obtained at a certain step the so-called operations of **conditional transfer of control** (conditional jump) are used.

The examples below illustrate this kind of operation.

(1) *Transfer of control depending on relative value of two numbers taken with their algebraic signs.* This is an instruction of the form

$\lessgtr$ jump	α	β	k

When reached in the course of a program it causes the computer to compare the numbers stored in locations α and β. If the former is greater than the latter the computer executes the next instruction in the original sequence, if otherwise it goes to the instruction stored in location k.

(2) *Transfer of control depending on relative magnitude of two numbers.* We denote this instruction by

$\lvert\lessgtr\rvert$ jump	α	β	k

and understand it in the sense of the previous operation with the only difference that it is the absolute values of the numbers which must be compared.

(3) *Transfer of control depending on the sign of a number (pl jump).*

Plus jump	α	k_1	k_2

This operation means: "If the number in location α is positi execute the instruction stored in location k_1, otherwise execute tl instruction stored in location k_2."

The latter instruction may be used for the operation of uncond tional transfer of control to location k. For this purpose it is sufficie to form the following instruction:

Plus jump	α	k	k

This instruction transfers control to the instruction stored in loc tion k irrespective of the number stored in location α.

A computer performs any of the available operations after tl corresponding instruction has been received. The instructions a represented by binary numbers and stored in computer's memo as well as initial data. The control unit interprets these instru tions, that is decodes them and applies the proper signals to tl other parts of the machine.

7. **Realization of Operations Within a Computer.** From the stan point of engineering every process of computing consists in co verting the corresponding pulses by electronic circuits. We a not going to discuss this question at length, that is to go into part culars concerning the elements needed for performing specific op rations. We shall only limit ourselves to the operation of additi of two positive quantities by way of illustration.

Operation of addition in the binary system, as in any other pos tional number system, involves the bit-to-bit addition and, wh necessary, a carry to the next higher digit place. The bit-to-b addition complies with the following rules:

$$0+0=0; \qquad 1+0=0+1=1 \quad \text{and} \quad 1+1=0$$

plus next place unity.

Let a and b be the digits we have to add together when performi the operation of addition on a certain digit place and let c be tl carry from the foregoing place. To perform the operation of add tion on a digit place means the following: given a, b, and c whi may take one of the two values zero and unity, it is required find s (the digit to be written into this place) and p (the carry the next place). The variety of all the situations that may ari here is confined to the following table:

a	0	1	0	0	1	1	0	1
b	0	0	1	0	1	0	1	1
c	0	0	0	1	0	1	1	1
s	0	1	1	1	0	0	0	1
p	0	0	0	0	1	1	1	1

It follows that to realize the addition on one digit place the machine must have a device with three values (a, b, and c) at its input and two values (s and p) at its output. The operation of this device must be in accordance with the above table. Therefore, when voltages at the inputs are equal to zero the voltages at the outputs do not exist either; when voltage is fed to one of the inputs the output voltage is produced at s while that at p is equal to zero, and so on.

Such a device is referred to as an **adder**. It can be easily realized as a circuit composed of electronic valves or transistors but we shall not discuss this question in detail here.

§ 3. ELEMENTS OF PROGRAMMING

1. General Notions. A problem to be solved on a computer must be expressed as a sequence of elementary operations which the machine is able to perform. Each operation is specified by the corresponding instruction, and the complete sequence of instructions necessary to solve the problem is referred to as the **program or routine. Programming**, that is preparing a routine, is one of the main stages of computing. It is clear that prior to this stage an appropriate mathematical method for solving the problem must be chosen and specific formulas for computing should be obtained.

The form of a routine depends on the numerical method chosen for solving the problem (for instance, we may compute an integral by means of the trapezoid rule, rectangular rule or some other formula of approximate integration) and on the type of the machine, i.e. on the set of operations it can perform. When the numerical method and the type of the machine are fixed the routine is not yet uniquely determined because there is a variety of ways for reducing the calculations to elementary operations. The selection of the most suitable routine depends, to a considerable extent, on

the skill of the programmer. Here we shall not go into particulars and confine ourselves to some elementary examples.

2. Formula Programming. Problems involving formulas are the simplest for solving on a computer. In this case programming is reduced to representing the formula as a sequence of elementary operations in a reasonable way and to placing the corresponding instructions and initial data into the storage. Let us consider a simple example.

Example. Given x, it is required to evaluate the quantity y expressed by the formula

$$y=\frac{2x+3}{5x+1}$$

The calculations can be represented as the following sequence of elementary steps:

$$(1)\ A_1=2x, \quad (2)\ A_2=A_1+3, \quad (3)\ B_1=5x,$$
$$(4)\ B_2=B_1+1, \quad (5)\ y=\frac{A_2}{B_2} \tag{1}$$

To perform these operations on a computer we arrange initial data in five storage locations (e.g. locations with numbers from $n+1$ to $n+5$). This can be written down as

Location	Stored number	Location	Stored number
$n+1$	x	$n+4$	5
$n+2$	2	$n+5$	1
$n+3$	3		

Let us place the instructions corresponding to the above operations into some other five locations. We thus obtain the following sequence of instructions:

Location	Operation	1st address	2nd address	3rd address	Result of operation
$m+1$	Multiplication	$n+1$	$n+2$	$n+2$	$2x$
$m+2$	Addition	$n+2$	$n+3$	$n+2$	$2x+3$
$m+3$	Multiplication	$n+1$	$n+4$	$n+1$	$5x$
$m+4$	Addition	$n+1$	$n+5$	$n+1$	$5x+1$
$m+5$	Division	$n+2$	$n+1$	$n+2$	$\frac{2x+3}{5x+1}$

When an intermediate result is not needed for further computations it may be erased so that the storage location can be used for new data. For instance, the execution of the first instruction involves such a procedure: the product of the numbers stored in locations $n+1$ and $n+2$ is again placed into location $n+2$. This makes the use of the storage more effective without overloading it with unnecessary data.

The program we have composed must be further supplemented with an input instruction which transfers the initial data and instruction codes into the storage and with the instruction for decimal-to-binary conversion because the machine uses the binary system while initial data are usually written and fed to the machine in the decimal notation. Instruction "division" written in location $m+5$ must be followed by three more instructions, namely, for converting the result of computation to the decimal system, printing the answer and stopping the machine.

The final stage of preparing a program involves replacing symbolic addresses by the corresponding concrete numbers.

These numbers, the absolute addresses, are four-digit numbers which are written in the octal number system beginning with 0000. Several initial storage locations are commonly used as *standard operating cells* (for input operations, number system conversion etc.). Suppose we have a machine in which the first eleven locations are used for this purpose (locations with numbers from 0001 to 0013 in the octal notation)*. Then we begin the storage allocation with location 0014 and thus write down our program in the following form:

Location	Operation or number	1st address	2nd address	3rd address
0014	Input (in locations 0015 through 0032)	0015	0015	
0015	Conversion from decimal system to binary (locations 0026-0032)	0026	0004	0026
0016	Multiplication	0026	0027	0027
0017	Addition	0027	0030	0027
0020	Multiplication	0026	0031	0026
0021	Addition	0026	0032	0026
0022	Division	0027	0026	0027
0023	Conversion from binary system to decimal	0027		0027
0024	Print	0027		

* Let the location identified by 0000 contain the number zero.

Continued

Location	Operation or number	1st address	2nd address	3rd address
0025	Stop			
0026	x			
0027	2			
0030	3			
0031	5			
0032	1			
0033				

3. **Cyclic Processes.** It is clear that in such a simple case as we have just considered there is no practical sense in using a computer. This elementary example has been taken with the only purpose to demonstrate how some familiar expressions are translated to the language intelligible to the machine. The application of a digital computer is effective only when the number of operations that the machine carries out is much greater than the number of instructions we must feed into its memory. An instruction can be reused many times when a repetitious series of the same operations is present in a problem. In such cases routines are said to contain **loops**. The action of performing each operation in one traversal of a loop is called a **loop cycle**. We now consider two elementary examples of cyclic programs.

(1) *Computation of the Square Root.* Suppose it is necessary to approximate the square root of a, a positive number, to a given accuracy. For solving this problem we may take advantage of the following fact (e.g. see [8], Chapter 3). The sequence

$$x_0 = a, \quad x_1 = \frac{1}{2}\left(x_0 + \frac{a}{x_0}\right), \quad x_2 = \frac{1}{2}\left(x_1 + \frac{a}{x_1}\right), \quad \ldots$$
$$\ldots, \; x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), \; \ldots \tag{2}$$

converges to $\sqrt{a}$ for every positive number a. Calculating the successive approximations x_1, x_2, ... and so on we may continue the process until a sufficiently accurate value is obtained. For testing the accuracy of an approximation we can compare x_n with x_{n-1}; when the difference is less than the preassigned value we stop the iterative process.

Now we see that to compute $\sqrt{a}$ we must place into three storage locations the three numbers: $a = x_0$ (assumed to be an initial approximation), ε (assessing the degree of accuracy) and $\frac{1}{2}$. The further computation of $\sqrt{a}$ is performed according to the following routine:

Location	Operation or number	1st address	2nd address	3rd address	Description of operations and their results
0014	Input	0016	0015		Inputting numbers and instruction codes
0015	Decimal-to-binary conversion	0030	0003	0030	Converting input data to binary system
0016	Bit-to-bit addition	0031		0027	Sending x_n from location 0031 (where x_n remains to be stored) to location 0027
0017	Division	0030	0027	0031	$\frac{a}{x_n}$
0020	Addition	0031	0027	0031	$x_n+\frac{a}{x_n}$
0021	Multiplication	0031	0033	0031	$x_{n+1}=\frac{1}{2}\left(x_n+\frac{a}{x_n}\right)$
0022	Subtraction	0031	0027	0027	$x_{n+1}-x_n$
0023	$\vert \lessgtr \vert$ jump	0032	0027	0016	Checking whether the given accuracy is achieved by comparing $\vert x_{n-1}-x_n\vert$ with ε and completing the cycle if the accuracy is attained
0024	Binary-to-decimal conversion	0031		0031	Converting the result to decimal system
0025	Print	0031			Printing the result
0026	Stop				
0027					Operating cell
0030	a				
0031	x_0				
0032	ε				
0033	$\frac{1}{2}$				
0034					

(2) *Tabulating Functions*. Another typical example of a cyclic process is evaluating elementary functions for different values of their arguments, i.e. compiling tables.

For example, let us consider the function $\sin x$. From Taylor's formula it follows that

$$\sin x = x-\frac{x^3}{3!}+\ldots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+R_{n+1} \qquad (3)$$

where the remainder R_{n+1} satisfies the inequality

$$|R_{n+1}| \leqslant \frac{x^{2n+3}}{(2n+3)!}$$

Denoting the kth term of the sum on the right-hand side of (3) by u_k and setting $s_k = u_1 + u_2 + \ldots + u_k$ we get

$$u_{k+1} = -x^2 \frac{u_k}{a_k}, \quad \text{where } a_k = 2k(2k+1) \tag{4}$$

and

$$s_{k+1} = s_k + u_{k+1} \quad (k = 1, 2, \ldots), \quad s_1 = x \tag{5}$$

Finally, it is readily seen that

$$a_{k+1} = a_k + 8k + 6, \quad a_1 = 6 \quad (k = 1, 2, \ldots) \tag{6}$$

We thus arrive at the following computational procedure: we choose a rough approximation $s_1 = x$ for $\sin x$ and compute the successive approximations s_k ($k = 1, 2, \ldots$) by using formulas (5).

At each step, after the kth approximation s_k has been computed, we first find a_{k+1} (by (6)), then u_{k+1} (by (4)) and finally s_{k+1}. If u_{k+1} appears to be less than the preassigned number ε the machine regards s_k as the sought-for value of $\sin x$ and proceeds to compute $\sin x$ for another x. This computational scheme may be realized by means of the following program:

Location	Operation or number	1st address	2nd address	3rd address	Description of operations and their results
0014	Input	0043	0015		Inputting numbers and instruction codes
0015	Decimal-to-binary conversion	0044	0010	0044	Converting input data to binary system
0016	Bit-to-bit addition	0050		0062	Transferring x to the standard location for u_k
0017	Bit-to-bit addition	0062		0063	Transferring u_1 to the standard location for s
0020	Multiplication	0062	0062	0064	x^2
0021	Subtraction		0064	0064	$-x^2$
0022	Multiplication	0045	0060	0065	$8(k-1)$
0023	Addition	0065	0046	0065	$8(k-1)+6$
0024	Addition	0057	0065	0057	a_k
0025	Addition	0060	0044	0060	k
0026	Division	0062	0057	0065	$\frac{u_k}{a_k}$

Continued

Location	Operation or number	1st address	2nd address	3rd address	Description of operations and their results
0027	Miltiplication	0065	0064	0062	$-x^2 \frac{u_k}{a_k} = u_{k+1}$
0030	Addition	0063	0062	0063	$s_{k+1} = s_k + u_{k+1}$
0031	$\|\leqslant\|$ jump	0047	0062	0022	Completing computation of sin x_i
0032	Bit-to-bit addition	0063		0050	Transfer: sin $x_i \to x_i$
0033	Bit-to-bit addition			0060	Transfer: $0 \to k$
0034	Bit-to-bit addition			0057	Transfer: $0 \to a_0$
0035	Address modification	0016	0056	0016	Modifying the 1st address of the instruction in location 0016
0036	Address modification	0032	0055	0032	Modifying the 3rd address of the instruction in location 0032: $i \to i+1$
0037	Addition	0061	0044	0061	
0040	Transfer of control	0061	0054	0016	Completing tabulation
0041	Binary-to-decimal conversion	0050	0003	0050	
0042	Print	0050	0003		
0043	Stop				
0044	1				
0045	8				
0046	6				
0047	ε				
0050	x_1				
0051	x_2				
0052	x_3				
0053	x_4				
0054	4				
0055	1 in the 3rd address				
0056	1 in the 1st address				
0057	$0\ (a_k)$				
0060	$0\ (k)$				
0061	Operating cell for i				
0062	Standard location for u_k				
0063	Standard location for s				
0064	Standard location for $-x^2$				
0065					

Similar routines can be written for evaluating the other elementary functions ($\cos x$, e^x, $\ln x$ etc.).

4. Flow-chart. Subroutines. When preparing a program for solving a complex problem it is helpful to represent it as a sequence of blocks corresponding to individual problems, that is to draw a **flow-chart.** This makes programming easier; moreover, one and the same block may be involved in different routines as a **subroutine.** Let us consider an elementary example. Suppose we have to compute an approximate value of an integral

$$J = \int_a^b f(x)\,dx \tag{7}$$

by using the rectangular formula (e.g. see [8], Chapter 12). It is natural to break down the computation into two steps (two blocks):

(1) Computing the values of the function $f(x)$ for the points x_i involved in the rectangular formula.

(2) Computing the sum

$$S = \frac{b-a}{n} \sum_{i=1}^{n} f(x_i) \tag{8}$$

which approximates integral (7). The routine for computing $f(x_i)$ depends upon the form of $f(x)$, while that for evaluating sum (8) is irrelevant of the choice of $f(x)$.

The absolute error of the result (i.e. the absolute value of the difference $J - S$) depends on two factors: the accuracy of the rectangular formula* and the accuracy with which the values of f at the points x_i are determined.

5. Instruction Codes. Operations on Instructions. In writing the above programs we referred to machine operations by symbolic names such as "addition", "multiplication" etc. These symbols, however, must be replaced by the corresponding binary numbers to be entered into the machine, i.e. by **instruction codes.** These binary numbers are fed to the memory locations in the same way as data, that is in the form of a sequence of the symbols 0 and 1. Each location has several predetermined binary places for storing them.

Let us consider an imaginary three-address computer. Suppose its locations are of 42-digit capacity, a 6-digit number being used for the operation code and three 12-digit numbers representing the addresses. Then the instruction "add the numbers in locations 7

---

* For estimation of errors of various formulas for approximate integration see [8], Chapter 2, § 2.

and 12 together and store the result in location 13" is represented in the code of our machine as

$$\underbrace{0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1}_{\text{1st address}}\,\underbrace{0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,0\,0}_{\text{2nd address}}\,\underbrace{0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,0\,1}_{\text{3rd address}}\,\underbrace{0\,0\,0\,0\,0\,1}_{\text{Operation code}}$$

where 000001 is assumed to designate the operation of addition. The fact that instructions and numerical data, when put into the machine, are of similar form presents no difficulties. Moreover, this enables us to treat instructions as ordinary numbers, e.g. to add* one instruction to another, which turns out to simplify programming. To illustrate this let us consider a simple example. Suppose we have to determine the sum of a thousand numbers. We may of course place them in the storage, e.g. in the locations with numbers from $n+1$ to $n+1000$, and then compile the following program:

1st instruction:

Addition	$n+1$	$n+2$	$n+2$

2nd instruction:

Addition	$n+2$	$n+3$	$n+3$

. . . .

999th instruction:

Addition	$n+999$	$n+1000$	$n+1000$

But we can solve this problem more economically: we again place the numbers in locations $n+1$ through $n+1000$ and then write (e.g. in location $n+1001$) the following "instruction":

	0001	0001	0001

Now let an instruction of the form

Addition	$n+1$	$n+2$	$n+2$

---

* It should be noted that operations on instructions involve special types of addition, namely, operation code addition, address-to-address addition and bit-to-bit addition.

be stored in location $m+1$. This instruction results in adding together the first two numbers. In the next location $m+2$ we place the following instruction:

Address-to-address addition	$m+1$	$n+1001$	$m+1$

In location $m+3$ we put the instruction of transferring control to location $m+1$ which will at this stage contain an instruction of the form:

Addition	$n+2$	$n+3$	$n+3$

The latter instruction causes the machine to add the third number to the sum of the first two. It is clear that this loop of three instructions will provide adding together all the numbers stored in locations $n+1$ through $n+1000$. It remains to provide the corresponding instructions for printing the answer and stopping the machine when the work is completed.

Thus, the use of operation of addition of instructions has made it possible to perform **address modification** and thus replace a long chain of similar instructions by a small number of operations.

6. Automatic Programming. Modern computer techniques include libraries of subroutines and other programming facilities but nevertheless the procedure of programming takes much time and effort. The stage of programming may require much more time than that of machine operation. This primarily applies to modern large high speed computers, and particular attention is therefore attached to various methods of *automatic programming*.

Here we cannot go into particulars concerning these methods. Their basic idea is to use the computer for translating a routine written in a symbolic language into a machine language. In other words, a mathematician describes the procedure for solving a problem in a special language using words and symbols taken from a fixed set of terms (vocabulary). Each character in this text when transferred into a machine is represented by a definite combination of zeros and ones. A special routine called the **translator** is then used to translate this program into a machine language program that can be run on a computer. It is necessary that the procedure be described in accordance with some formal rules of syntax by means of a clear-cut set of available words. There exist several conventional languages designed for automatic programming. The most widely used languages are ALGOL and FORTRAN*. Each

* Abbreviations for "Algorithmic Language" and "Formula Translating".

language can be used independently of the computer at hand while the choice of the translator is governed by the choice of a language and the type of a machine (but not by the problem to be solved).

The use of algorithmic languages and translators reduces tedious and labour-consuming work of a programmer.

§ 4. ORGANIZATION OF COMPUTER WORK

1. Conditions for Effective Use of a Computer. As has been said, the solution of a problem on a computer requires an adequate routine, i.e. a sequence of coded instructions causing the computer to perform the elementary operations which the problem is reduced to. If the number of individual instructions in a routine were equal to the number of operations necessary to solve the problem, the use of computers would not be effective since the process of programming would take almost as much time as performing all the calculations manually. However, every complicated problem which is to be solved on a computer contains specific groups of operations that must be repeated over and over again. This was shown in the instance of the routine for computing the square root; to a greater extent this is associated with more complex problems. The number of instructions in a reasonable routine is therefore much less than the number of operations it causes the computer to perform. The use of a computer is particularly effective in the problems where different sets of initial data must be processed in one and the same manner. On the other hand, there exist problems in which the application of a computer is useless because they require a long and tedious process of programming while the amount of calculations is comparatively small.

The question whether the computer methods are suited for a particular problem is of primary importance for effective use of computers.

2. Basic Stages of Solving a Problem on a Computer. Once the decision to compute has been made, the problem moves through the following stages:

(1) *Problem Formulation.* In the first place any problem to be solved on a computer must be stated precisely in the mathematical form. In other words, the problem that has arisen in physics, engineering or elsewhere must be described as that of solving equations, evaluating integrals and so on. It should be stressed that this stage is mainly based on collaboration of physicists or engineers setting the problem and specialists in computational mathematics and is therefore connected with considerable difficulties. To overcome them successfully it is necessary that the mathematicians be familiar with the nature of the problem they deal with. On the other hand, the engineers and physicists must have a clear idea about compu-

tational methods and the advantages they put at his disposal.

(2) *Problem Analysis* (*Selection of an Algorithm*). Computer cannot deal with such terms as the solution of an equation, integral, function etc. in which we commonly state a problem. Hence, before translating the problem into a language acceptable to the machine, a suitable numerical method must be selected for such operations as evaluation of derivatives and integrals, solution of equations, etc. For example, derivatives are replaced by the corresponding divided finite differences, integrals are evaluated by means of some approximate formulas (such as the trapezoid rule or Simpson's rule) and so on. Thus, the problem is reduced to a finite sequence of arithmetical operations. It is clear that one and the same problem can be solved by using different numerical methods. Computer efficiency essentially depends upon the proper choice between alternative methods of computation.

(3) *Programming*. This stage comes when the problem analysis has been accomplished, i.e. when a suitable algorithm consisting of a sequence of elementary operations has been chosen for each step. The routine, however, may be written in a variety of ways. The choice of the most adequate program, of the best way of using the storage and other machine's facilities requires professional skill of the programmer, his experience and familiarity with the type of the machine.

(4) *Machine Run*. When a program has been completed the process of computations reduces to a standard machine operation. The person who manipulates the controls of a computer (operator) may be unfamiliar with the particulars of the problem he deals with.

3. Checking Computer Operation. Error Detection. Mass-scale calculation on a computer involves millions of elementary operations. It is therefore a complicated problem to prevent computer errors. Wrong results may be obtained for various reasons. To begin with, the program itself may have errors. A single mistake in a routine may result in the failure of the whole computation process. Hence, any mistakes in a computer program must be located and corrected before starting computations. This is the so-called *checkout process*. Sometimes it is helpful to perform a step of calculation manually and then compare the result with the machine result at the same step. There exist some other systematical methods of checking programs which however we shall not discuss here.

Another thing we must make sure of is the proper hardware operation. This is usually checked by putting through a number of test routines. Some of these are specially designed to test all the main units and functions of the computer. But it should be taken into account that a failure of equipment may occur in the process of calculation. Such faults are detected and eliminated in the following

manner: the machine is made to store an intermediate result and to repeat computation once again. If the two results coincide the machine proceeds to the next step.

Finally, it may be a rounding error that involves wrong results. The finite capacity of storage locations puts a limit to the precision of computations since all the quantities must be rounded off by deleting less significant digits. These deviations from the theoretically correct values may occur in the course of a long computation and yield an incorrect result without any fault in the routine or hardware.

There exist different methods for increasing the accuracy with which the computed results are obtained. For instance, we may write numbers having twice as many digits as are normally handled in a given computer by using two locations instead of one for each number (the so-called **double-precision** method).

BIBLIOGRAPHY

(Starred items are in Russian)

1.* Bari, N. K., *Trigonometric Series*. Fizmatgiz, Moscow, 1961.
2.* Berezin, I. S., Zhidkov, N. P., *Computational Methods*, vol. 2, Chapter 4, Fizmatgiz, 1962.
3. De Brujn, N. G., *Asymptotic Methods in Analysis*, Amsterdam, North-Holland Publishing, 1961.
4. Eddington, A. S., *The Mathematical Theory of Relativity*, Cambridge, Cambridge University Press, 1924.
5. Elsgolts, L. E., *Differential Equations and Calculus of Variations*, Mir Publishers, Moscow, 1970.
6. Erdély, A., *Asymptotic Expansions*, New York, Dover, 1956.
7.* Gelfand, I. M., Shilov, G. E., *Generalized Functions*, vols. 1, 2, 3, Fizmatgiz, 1958.
The English translation: *Generalized Functions*, Academic Press, Inc., New York, 1964.
8.* Ilyin, V. A., Poznyak, E. G., *Fundamentals of Mathematical Analysis*, "Nauka", Moscow, 1967.
9.* Ivanenko, D. D., Sokolov, A. A., *Classical Field Theory*, GTTI, Moscow, 1951.
10.* Kochin, N. E., *Vector Calculus and Elements of Tensor Calculus*, GONTI, Moscow, 1938.
11.* Kolmogorov, A. N., Fomin, S. V., *Elements of the Theory of Functions and Functional Analysis*, "Nauka", Moscow, 1968.
The English translation: *Introductory Real Analysis*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1970.
12.* Lavrentyev, M. A., Shabat, B. V., *Methods of the Theory of Functions of a Complex Variable*, "Nauka", Moscow, 1965.
13.* Norden, A. P., *Differential Geometry*, Uchpedgiz, Moscow, 1948, p. 186.
14.* Rashevsky, P. K., *Riemannian Geometry and Tensor Analysis*, "Nauka", Moscow, 1964.
15.* Sobolev, S. L., *Equations of Mathematical Physics*, GTTI, Moscow, 1954.
16.* Tikhonov, A. N., *Stable Methods for the Summation of Fourier Series*, Doklady Akademii Nauk SSSR, vol. 156, N 2, 1964, 268-271.
17.* Tikhonov, A. N., Samarsky, A. A., *Equations of Mathematical Physics*, "Nauka", Moscow, 1966.
The English translation: *Equations of Mathematical Physics*, New York, Pergamon Press, 1963.
18.* Yefimov, N. V., *Higher Geometry*, Fizmatgiz, Moscow, 1961.

NAME INDEX

SUBJECT INDEX

TO THE READER

Mir Publishers would be grateful for your comments on the content, translation and design of this book. We would also be pleased to receive any other suggestions you may wish to make.

Our address is:
USSR, 129820, Moscow I-110, GSP
Pervy Rizhsky Pereulok, 2
MIR PUBLISHERS

Printed in the Union of Soviet Socialist Republics